Unwinding Financial Market Complexity

Morales, Raffaello

Awarding institution:
King's College London

King's College London

Doctoral Thesis

Unwinding Financial Market Complexity

Author:

Raffaello Morales

Supervisor:

Tiziana Di Matteo

A thesis submitted in fulfilment of the requirements

for the degree of Doctor of Philosophy

in the

Financial Mathematics group

Department of Mathematics

May 2014

Declaration of Authorship

I, Raffaello Morales, declare that this thesis titled, 'Unwinding Financial Market Complexity' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: _____

Date: _____

"Zum Ziele führt dich diese Bahn, doch muss du Jüngling männlich siegen. Dann höre unser' lehre an: sei standhaft, duldsam und verschwiegen."

- E. Schikaneder -

"Science is organised knowledge. Wisdom is organised life."

- I. Kant -

"I don't see any problem in being faced with a problem. I see a problem in putting up to our own ineptitude to its solution".

- R.M. -

"Nemo reperimur, qui sit studio nihil consecutus".

- Quintiliano -

"Morality and values depend on the existence of conscious minds - and specifically on the fact that such minds can experience various forms of well-being and suffering in this universe."

- S.Harris -

Published Articles

I list here the peer reviewed articles I have produced in the course of my PhD and whose material I have used in this thesis. The articles are listed in chronological order of appearance:

1. R.Morales, T.Di Matteo, R.Gramatica and T.Aste, *Dynamical Generalized Hurst Exponent as a Tool to Monitor Unstable Periods in Financial Time Series*, Physica A, 391, 2012, 3180-3189.

2. R.Morales, T.Di Matteo and T.Aste, *Non-Stationary Multifractality in Stock Returns*, Physica A, 392, 2013, 6470-6483.

3. R.Morales, T.Di Matteo and T.Aste, *Dependency Structure and Scaling Properties of Financial Time Series are Related*, Scientific Reports, 4, 4589; DOI:10.1038/srep04589 (2014).

Results presented in Paper 1 are reported in Chapter 2 Section 2.5.
Results presented in Paper 2 are reported in Chapter 3.
The results reported in Paper 3 are presented in Chapter 4 and 6.

Abstract

Complex systems are characterised by different distinguishing aspects often associated with completely separate behaviours. In financial markets, paramount example of complex systems, two of these aspects stand out in characterising the statistical properties of the many constituents: one is multifractality, a feature which describes the departure of financial time series from purely random processes and is therefore a measure of complexity of the prices; the other is the cross-correlation structure between assets, which encloses information about the market organisation and can reveal dominant factors as well as hierarchical properties.

In this thesis I have studied the relationship between these two distinctive properties of financial markets. I have first unveiled new empirical properties of stock returns, casting new light on the latent mechanism governing price dynamics and interactions, and I have then proposed a model which reproduces the observed properties.

I have investigated multifractality dynamically on stock returns after having introduced the weighted generalised Hurst exponent, a study that has revealed remarkable increasing trends in the dynamical scaling exponents for firms bailed-out after the 2008 financial crisis. I have then tested the significance of dynamical fluctuations of multifractality against a well-established multifractal model, the Multifractal Random Walk (MRW). The hypothesis of constant multifractality in financial markets has been rejected in many cases revealing a much more complex behaviour of financial time series. I have then linked the multifractal behaviour in financial markets to the cross-correlation structure, showing that the two properties are indeed related. I have investigated the relationship between a proxy of multifractality and cross-correlation hierarchical properties on different markets which have confirmed the result. After having thoroughly reviewed the existing literature on multivariate models, I have proposed a dynamical multivariate model able to reproduce the empirical facts reported in this thesis along with an array of

other well-established stylised facts, thus unifying correlation and multifractality in a unique coherent framework.

Acknowledgements

It is a great pleasure to express my deep heartfelt gratitude to the many who have made this three year long doctoral experience unique. Sometimes I wonder how things would have turned out if even a single person contributing to my everyday life were not to be part of it.

London being such an overwhelming and ever changing setting, I had the chance and the luck to know and spend time with many people: of these people, someone has moved elsewhere, others have just arrived, and yet others really do not want to beat feet - still to acknowledge for their endurance.

My first thanks go to my principal thesis supervisor Tiziana as well as to my second supervisor Tomaso. Let alone their efforts in guiding me to produce some relevant science - which all but failed - they have contributed to make the overall research experience an extraordinary one, always slipping in a laugh at the right time and never taking themselves too seriously. Tiziana has proved to be much more than a good thesis advisor, constantly demonstrating her fondness for her students and always struggling to deliver the best for their growth. I really want to thank her for that. Tomaso on the other hand has been able to provide, alongside bracing brainstorming sessions, the necessary confidence in me that being fifty can still be fun after all. Looking forward to it! I also wish to thank Ruggero Gramatica for his helpful advices on career perspective as well as for his hints on research. Andrea Macrina is also warmly acknowledged for his advices.

For their mentoring role I wish to warmly thank Paul and Alison Madden, Andrew and Beth Taylor, Keith McEwan, Massimo Bianchi and Philippe Moschetta: having the opportunity to hear their opinion on crucial decisions has been invaluable.

I will then start to list the constellation of people who have enriched my life-experience over the last three years. Big thanks to Francesco for his constant mentoring help during my first months at KCL as well as for the fact of being one of the nicest persons I have ever come across in my life. His personal evolvement and consulting guide in some of the most tricky steps to take were invaluable. Thanks to him also for having introduced me to Ruggero, whose friendship is today more than what the word may evoke. Further to this, a specific extra thank to him is the following: thank you so much for not being vegetarian any longer!

I want to thank all the students at KCL for their success in daily coping with the strenuous task of disproving the equality PhD-student=nerd[1]: Daphne, Nicoló,

[1]Sorry for those who are not mentioned here: but never let it go, try harder.

Paolo, Michael, Matt, Giovanni and many more (?). Special thanks to Nicoló for being very supportive in the research, providing helpful advices and stimulating discussions.
I wish to thank Paolo for adding some genuine roman wit to my days, Silvia for all her constant support and fondness, Benedetta for playing the little sister so well - she almost let me not regret not having a real one. A handful of thanks then to the vibrant community of City Mansions for making my everyday life in these last two and half years an unforgettable experience: thanks to Carlos, Ana, Lewis, Matt, Lisa, Tim, Imke and special thanks to Michele, for being such a great flatmate other than a great artist and above all a much beloved friend: thanks Maestro!
Thanks to Cash for being such a special and unique friend and for his untiring efforts in turning me away from work. Our relationship is the best example of how open-mindedness, generosity and respect can bring together two utterly different men. Warm thanks to Andrea for our amazing friendship and for being always a good guy benchmark: he stands as one of the persons I most enjoy spending time with.
I would also like to thank all the people who have supported me from the distance and have visited me repeatedly in the course of this PhD: Jack, Pepa, Giulia, Ale, Sophie, Edo, Pupo, Giacomo - thanks for being the only friend not having visited me in almost 5 years - , Tony, Mathieu, Philipp, Andrea, Antonio, Simon, Mancio, Benji, Magro, Serri, Isotta, Antonio, Federico, Alessandra, Francesco, Pala and I'm sure I'm leaving someone aside.
I want to thank Carlotta for her constant support and love over these many years abroad and for having provided the necessary energy and spark when my motivations seemed to be wavering.
Stefano is also warmly acknowledged for providing great company and fun over these last months as well as for his insightful hints and career advices and for having found a challenging and pleasant engagement to my sleepless early mornings.
I wish to thank Adolfo for having been an irreplaceable master in teaching me so many things that a university or a business school - let alone a music conservatory - will never teach you. He has been able to help me develop the necessary connection between heart and brains by helping me digging deep into my personality: he is a great friend other than a great mentor.
Finally, I want to express my most sincere gratitude and love to my wonderful solid three pillars: Andrea, Mom and Dad. Thanks for having always prompted me to

be critical and to look at the world as a bottomless well of opportunities, thus whetting my curiosity and my enthusiasm to tackle every problem with the same energy. Also, thank you for having been effective in teaching me that structure and resolution are the best tools to solve any ambiguous problem. In particular, I am grateful to Andrea for having been - despite being younger than me - a great model for me, transmitting to me his stubbornness and determination: he is always a valuable reference for me as I hope to be one for him. Thanks to my parents for being contributing to my education and growth in the best complimentary way one could ever aim for. Thanks to my mother for being an example of generosity and for her inexhaustible energy and determination, particularly in difficult times: I am glad of having learned so much from her, above all that grit and wit in complex situations are invaluable assets in today's multifaceted world. Thanks to my father for having played the mentor role in the best way I could ever fathom, putting infinite efforts in ensuring the best for me. I vividly acknowledge him for his irreplaceable contribution toward shaping my personality and valuing certain aspects of life that are often overlooked.

Contents

List of Figures

List of Tables

Abbreviations

ACF	Autocorrelation function
cdf	cumulative distribution function
CHS	Common Hawkes Shock
CPS	Common Poisson Shock
DBHT	Deterministic Bubble Hierarchical Tree
GARCH	Generalized auto-regressive conditional heteroskedasticity
GHE	**G**eneralized **H**urst **E**xponent
HKSE	Hong Kong Stock Exchange
i.i.d.	indipendend identically distributed
LSE	London Stock Exchange
MDHM	Multivariate Dynamical Hierarchical Model
MF-DFA	Multifractal detrended fluctuation analysis
MGARCH	Multivariate GARCH
MMRW	Multivariate Multifractal Random Walk
MMSM	Markov-Switching Multifractal Model
MRW	Multifractal Random Walk
MSV	Multivariate Stochastic Volatility
NYSE	New York Stock Exchange
pdf	probability distribution function
wGHE	Weighted Generalized Hurst Exponent
rv	random variable
TSE	Tokyo Stock Exchange
tMRW	Multifractal Random Walk with Student-t residuals

Symbols

$C_r(\ell)$	ACF of returns
$C_{r^2}(\ell)$	ACF of square returns
$C_{\lvert r\rvert}(\ell)$	ACF of absolute returns
$C_{\ln r}(\ell)$	ACF of logarithmic returns
$C_\sigma(\ell)$	ACF of the volatility process σ_t
$\mathbb{E}(\cdot)$	expected value
f	probability density function
$F_<$	cumulative distribution function
$F_>$	complementary cumulative distribution function
H	Hurst exponent
$H(q)$	generalized Hurst exponent
K_m	Bernoulli random variable
ℓ	lag
$M_q(t,\tau)$	q-moments
$M_q^*(t,\tau)$	empirical q-moments
n	hierarchical order
$N_t^{(e)}$	point process (Poisson or Hawkes)
p_m	probability associated with the Bernoulli variable K_m
r_t	log-returns
P_t	price at time t
S_t	a generic stochastic process
w_t	market common volatility factor
X	a generic random variable

x	a realisation of X
$Y_{n,t}$	volatility hierarchical factor
$z_{j,t}$	volatility factor in common Hawkes process
$Z_{j,t}$	common shock process
α	tail exponent
β	ACF decay exponent
γ_i	hierarchical path for stock i
Γ_i	set of nodes affecting stock i
$\delta_\tau P_t$	increments of the price P_t
$\delta_\tau S_t$	increments of the process S_t
ϵ_t	residual
κ	Kurtosis
$\tilde{\kappa}$	Excess Kurtosis
ζ_q	scaling function
ζ_q^*	empirical scaling function
$\lambda^{(e)}$	intensity for events of type e
λ_t	Hawkes intensity
ν	Student-t degrees of freedom
ν_t	Hawkes kernel
ω_t	log-volatility
σ_t	volatility
$\mathbf{\Sigma}$	Covariance matrix
Σ_{ij}	Covariance matrix entries
τ	time scale

To my family,
Andrea, Carla and Piero

Chapter 1

Introduction

"Nobody accepts randomness in his own success, only in his failure."

- N.N.Taleb -

The science of complexity has made major leaps forward over the last twenty years in unveiling the fine structure of many systems, from biology and neuroscience to social and economic systems, from traffic jams to earthquakes and from population dynamics to finance. The understanding of financial markets, in particular, has prominently benefited from the use of complex systems tools and a new discipline called Econophysics [1–3] has thrived over the last twenty years. One aim of this research area is understanding the underlying forces governing financial markets and finding realistic models able to capture the observed behaviour [4, 5]. The challenging part of this task lies in the by now ascertained fact that markets are, to say it with Benoit Mandelbrot, misbehaved [6], that is they present irregularities that clash with the Gaussian picture [7, 8], originally proposed to describe price dynamics. The presence of fluctuations several order of magnitude more frequent than those accounted for by the Gaussian statistics and their impact on financial markets and ultimately on the society have been famously yet wittingly introduced to the large audience by N.N.Taleb [9, 10]. Although sometimes difficult to describe statistically, the irregular behaviour allows to distinguish prices from

completely random processes and can therefore be taken as a clue of the mechanism behind price swings and can also serve as a warning for financial crashes and crises [11, 12].

The main motivation of this doctoral thesis is to further unveil hidden behaviours of the stock markets, at the same time trying to understand the coherence behind different aspects of prices complexity. For this reason we have concentrated on two of the most relevant properties of stock prices: multi-scaling (or multifractal) behaviour [13] and cross-correlation [14]. These two aspects reflect very different properties of the prices: multifractality is a property of the single time series and reflects the irregularities of the process across different time scales; cross-correlation quantifies how each stock depends on the others and therefore represents a measure of the market interconnection and interdependence. Can these two properties be somehow related and is it possible to devise a mechanism able to explain both of them coherently starting from the existing complex systems tools? In this thesis we tackled these questions, looking for the first time at the interplay between multifractality and correlation. After having confirmed the interdependence between the two quantities, we have proposed a simple model able to account at the same time for time-varying multifractality and dependence between hierarchical order and multifractality.

The thesis is organised as follows: in Chapter 2 we first discuss the most relevant financial markets empirical properties, remarking how the prices irregularities fall away as the return time scale is decreased. The study of prices behaviour across different time scales naturally introduces the concept of multifractality, which is discussed along with some empirical examples. We further highlight the different contributions to the observed multifractal behaviour and show, in agreement with previous research [15–17], how the distribution tails contribute the most to the degree of multifractality. We then studied the dynamical evolution of scaling exponents. In order to do so we introduce the *weighted* generalised Hurst exponent [18], a variant of the well-established generalised Hurst exponent, which incorporates exponential smoothing in the estimation in order to damp the effect of

fluctuations far in the past on recent observations. We show that scaling exponents serve as a good tool to track the evolution of the share price of companies bailed-out during the 2008 financial crisis.

Chapter 3 is devoted to validate the dynamical fluctuations of multifractality against a multifractal model envisaging constant scaling exponents [19]. The main goal is to assess which portion of multifractality fluctuations has to be ascribed to estimation noise, and how much is instead significant in identifying scaling breakdown or switches. We have performed the hypothesis test extensively on NYSE daily data, considering as a benchmark model the Multifractal Random Walk [20], probably the most popular multifractal model. We have found a large portion of stocks failing the constant multifractality test, indicating an even more complex behaviour of prices than that predicted by the multifractal model [19]. The analysis has been also repeated considering as benchmark models two modifications of the Multifractal Random Walk: the first one replaces Gaussian with t-Student distributed residuals, thus accounting for fatter tails; the second considers log-gamma volatility instead of the usual log-normal one, again introducing large fluctuations in the volatility. In both cases the number of stocks failing the test is drastically reduced, but still significant fluctuations in multifractality cannot be completely ruled out. The results presented in this chapter are very relevant for future modelling guidelines, as they disclose the non-stationarity of financial time series from a new perspective.

Chapter 4 deals with the main empirical result of this thesis [21], that is the interdependence between cross-correlation hierarchical order and degree of multifractality. Using the Deterministic Bubble Hierarchical Tree (DBHT) clustering algorithm [22], a deterministic method which does not depend on any thresholding parameters, we are able to build a hierarchy of the filtered correlation matrix, which is then used to associate with each stock its corresponding hierarchical order [21]. We then find a positive - although clearly non-linear - dependence between hierarchical order and degree of multifractality in NYSE data. This pattern is observed on the entire market, considering average multifractality observed for each fixed hierarchical order, as well as locally on specific clusters and market sectors.

Further robustness tests by means of bootstrapping are performed in order to ascertain the validity of the result. We have also investigated the same behaviour on other international stock markets including London Stock Exchange and Tokyo Stock Exchange, confirming what observed on NYSE. We have then shown that interdependence between cross-correlation hierarchical order and multifractality is also detected in time, with both quantities increasing before the unfolding of the crisis, thus reflecting an overall coalescence of the market which goes along with a pronounced increase in multifractality.

In Chapter 5 we review multivariate models. We first discuss static multivariate models [23], pointing out their lack of any description of stock returns temporal persistence. This shortcoming is fixed by Multivariate GARCH models and Stochastic Volatility models, both analysed in their details throughout this chapter. We then discuss how the introduction of factor models can reduce the number of independent random components, thus providing more parsimonious models with smaller number of parameters. Among stochastic volatility models we present the Multivariate Multifractal Random Walk [24] and its simplified version with one volatility factor only, which will serve as the basis for the model developed in Chapter 6. As a first step towards the inclusion of a hierarchical structure in the multivariate modelling, we investigate the use of Common Shock approach to model stock returns, with the aid of Hawkes process [25] to model volatility bursts.

Finally, in Chapter 6 we introduce and discuss a multivariate dynamical model [21], generalising multivariate stochastic volatility models in such a way to include the empirical observations reported in previous chapters of the thesis. In particular, our model is the first to include at the same time a hierarchical structure together with a non-trivial temporal dependency structure in the volatility. This is obtained by introducing a hierarchical term in the volatility, factorized out of a common volatility factor and disappearing for certain values of the hierarchy parameters. The model can thus be seen as a generalisation of a common factor stochastic volatility model - like those discussed in Chapter 5 - with the hierarchical factor acting as a perturbation, reflecting a finer structure of the volatility.

We show in detail the properties of the model, with particular attention to the parameters limits and demonstrate with numerical simulation its ability to capture the interdependence between multifractality and hierarchical order.

In the Appendices we report extra material and additional details to the content of the various chapters.

At the beginning of each chapter we give a short abstract summarising the material and main results of each chapter. Likewise, a conclusive summary is given at the end of each chapter.

Chapter 2

Empirical Properties of Financial Time Series

"Unless you have confidence in the ruler's reliability, if you use a ruler to measure a table, then you may also be using the table to measure the ruler."

- Wittgenstein -

In this chapter we review the most relevant empirical properties of financial time series [1, 26], with particular attention to the asset class of stocks. We discuss fat-tails, autocorrelation of returns and volatility [27] and multifractality [28], the latter standing out as the most universal property able to capture both time-correlations and fat-tails of stock returns distributions. We also perform a new analysis aimed at identifying the different components of multifractality in stock returns, adding confirmation to results already established in previous works [15, 16]. After reviewing the literature on the generalised Hurst exponent method, we introduce the weighted generalised Hurst exponent and investigate its dynamical evolution for firms bailed-out after the 2008 crisis. Parts of the results and analysis presented in this chapter have been published in the paper *"Dynamical Generalized Hurst Exponent as a Tool to Monitor Unstable Periods in Financial Time Series"* [18].

2.1 Introduction

Since the early 1900, economists and statisticians have been interested in understanding the time evolution of stock prices. The first and most notable example is that of Louis Bachelier who, in his PhD thesis titled *"Theorie de la Speculation"* [29], proposed to use, borrowing it from physics, the newly discovered formalism of Brownian motion to describe the dynamics of the *Paris Bourse* index. The hypothesis of Gaussian noise though was soon understood to be all but realistic in capturing the empirical behaviour of price changes and a plethora of modifications and refinements of the original random walk hypothesis has been investigated in depth [30–32]. The most ubiquitous empirical properties, also known as stylised facts, of daily price increments are the following [3, 26, 33]:

- Price changes show distributions such that the probability of observing events many standard deviations away from the mean is much larger than that envisaged by a Gaussian distribution: one refers to this property of the distribution as *fat-tails* [34].

- Price changes are far from being independent identically distributed (i.i.d.) random variables, but show highly persistent patterns, as a consequence of the memory present in the market [27, 35, 36]. The most notable effect arising from this fact is the so-called *volatility clustering*, which reflects the intuitive idea that a large price change today will necessarily feed back on price changes in the coming days and that the time needed for this memory to be lost is long enough to rule out even an exponential decay, but requires a power-law description.

- The previous two universal properties motivated the fundamental work of Benoit Mandelbrot (see the very recent account in [37] for a summary of his scientific achievements), who first noted that the same properties were captured by another very relevant empirical property of financial time series: multifractality [28, 32]. A multifractal process has remarkable complex properties which reflect the non-linear relationship between the distribution

of the process increments across different time scales. The multifractal properties of stock prices represent a valuable tool to asses the heterogeneity of the different contributions to stock returns volatility. Since the pioneering works of Mandelbrot, multifractality has been progressively investigated on several asset classes [13, 38, 39] and has become a downright stylised fact of financial time series. It is likewise ascertained that the multifractal behaviour is a consequence of the other stylised facts and that these contribute unequally to its presence in financial time series: in particular, there is evidence that the broad distribution of returns contributes to multifractaltiy more than long-range dependence [15–17, 40, 41].

The properties listed above are typical of price changes at relatively high frequency (from few hours to few days) [27]. When the scale of the price increment is made coarser, the irregularities tend to disappear and the Gaussian uncorrelated picture becomes more realistic.
Asset returns present also a variety of other stylised facts: leverage effect [42], correlations of extreme returns [43–45], volume/volatility correlations [46, 47], conditional heavy tails, asymmetry in positive and negative fluctuations. In the course of this chapter though we are going to concentrate on those listed above, performing empirical analysis on original sets of data. We show stock return empirical properties across different time scales, which naturally leads to introduce the concept of multifractality, also reviewed in detail along with the most important technique to measure it: the generalised Hurst exponent.
The original contributions of this chapter are the following

- We perform, by means of the generalised Hurst exponent method, an empirical analysis of a set of daily stock prices aimed at identifying the different contributions to multifractality.

- We introduce the weighted generalised Hurst exponent, a novel estimator for the scaling exponent which weights observations in order to reduce the effect of noise distant in the past on more recent observations [18].

- We study dynamically the weighted generalised Hurst exponent on a set of firms bailed-out as a consequence of the 2008 financial crisis [18].

The structure of this chapter is as follows: in Section 2.2 we discuss the two most relevant stylised facts of log-returns and show on empirical data how these empirical properties vary across different time scales, pointing out that the irregularities of prices gradually wind down as the frequency of observation is decreased. We then introduce multifractality in Section 2.3 along with the generalised Hurst exponent, the main estimation tool that will be used throughout the entire thesis. We also discuss the literature establishing the economic relevance of the generalised Hurst exponent in identifying the market stage of development. In Section 2.4 we then show on our dataset the influence of different stylised facts on the degree of multifractality. We then present in Section 2.5 the main results of this chapter: after introducing the weighted generalised Hurst exponent, we study its dynamical behaviour and discuss how its fluctuations in time are related to the tail exponents.

2.2 Main Properties of Price Changes

In this section we review the empirical properties mentioned in the previous section. We first introduce the main variable of interest - namely the log-returns - and motivate the investigation of empirical properties across different time scales. The different data sets analysed are introduced in the corresponding sections.

2.2.1 Why log-returns?

The very first step which has to be taken in any scientific analysis is to properly define the variables of interest. In quantitative finance, it turns out that the price, although fundamental for its economic meaning, is not actually the best variable to be studied statistically [1][1]. In order to clarify this, we report in the

[1]In the course of this thesis we will refer to the price of a generic asset at time t as P_t.

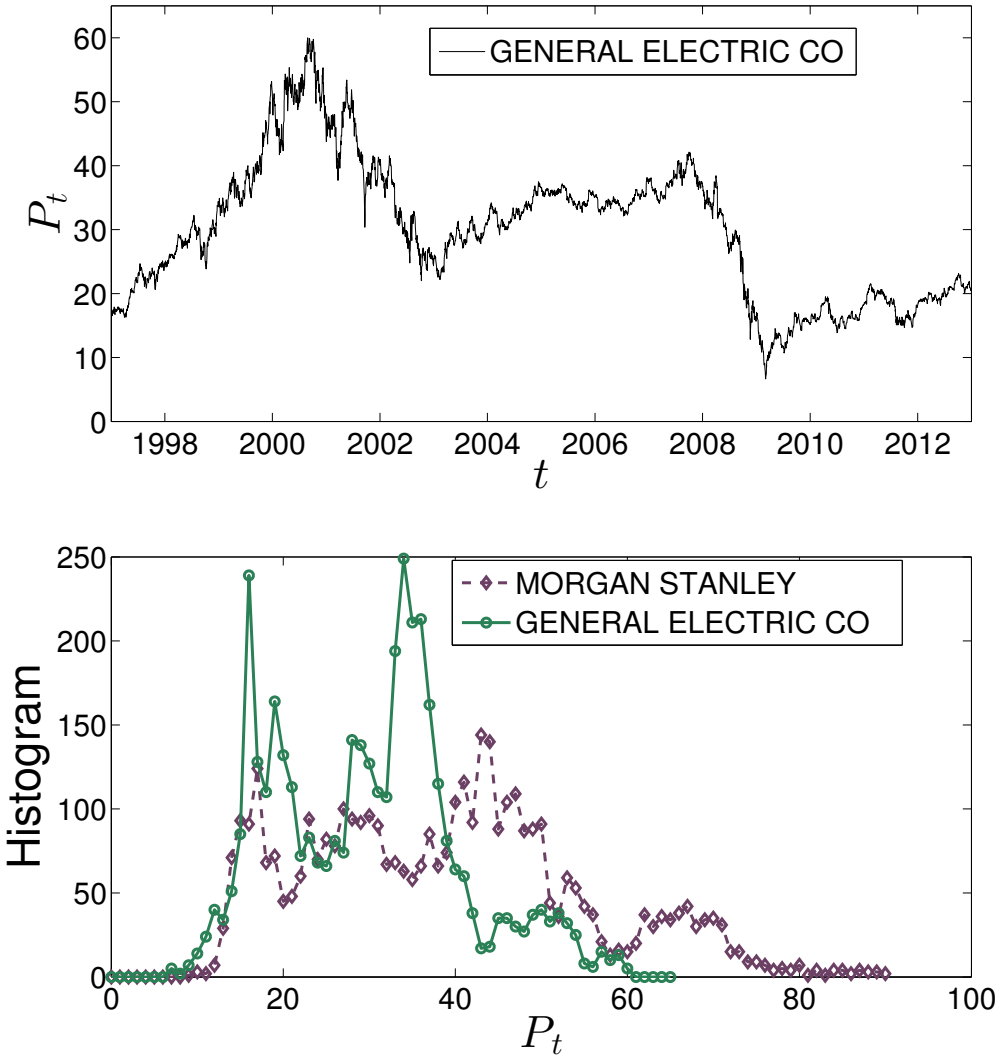

FIGURE 2.1: (Top) Daily price evolution of GE shares in the period 01/1997-12/2012. (Bottom) Histograms (elementary binning of one unity) of GE and MS daily prices in the same period.

top plot of Figure 2.1 the daily price of General Electric (GE) shares over fifteen years (in the period 01/1997-12/2012 [2]). In the bottom plot of Figure 2.1 we also show the histograms of the price values for both GE and Morgan Stanley (MS) shares. This last plot immediately tells that both prices show very irregular distributions, in spite of the pretty big size of the sample (4026 realisations). This is because the realisations of prices time series cannot be regarded as independent identically distributed (i.i.d.), but present very long range correlations. One can define several variables avoiding these problems and the most widely accepted in

[2]Data provided by Bloomberg.

TABLE 2.1: In this table we report the kurtosis of S&P500 and DJIA indices for daily, weekly and monthly returns. We note how the kurtosis decreases, resembling the expected value for a Gaussian distribution as we decrease the time scale at which fluctuations are observed.

	Kurtosis		
	Daily	Weekly	Monthly
S&P 500	10.2632	7.6747	3.9139
DJIA	10.2524	8.0636	3.0797

the modern financial literature is the log-return, defined, at scale τ, as [1, 2]

$$r_{t,\tau} = \log P_{t+\tau} - \log P_t. \tag{2.1}$$

Log-returns (or more simply returns) behave very similarly to the simple relative price increments

$$\delta_\tau P_t = \frac{P_{t+\tau} - P_t}{P_t}. \tag{2.2}$$

If one chooses $\tau = 1$ day (as will be done in most of this thesis) then the typical price change for stocks is of the order of a few percent, which entails a difference between $r_{t,\tau}$ and $\delta_\tau P_t$ of the order 10^{-3}, absolutely negligible for the purpose of the statistical analysis. We will then stick to returns as the fundamental variable of interest and we will denote simply by r_t daily returns.

2.2.2 One scale, many scales

An aspect which is central to this work, and that will be explored throughout the different chapters, is how returns behave at different fluctuation scales τ and how they scale with τ. In Table 2.1 we report values of the Kurtosis of the daily, weekly and monthly returns (corresponding to $\tau = 1, 5, 25$ respectively) for the two major American composite indices Stadard & Poor's 500 (S&P500) and Dow Jones Industrial Average (DJIA)[3] in the period 12/1999-12/2011[4]. We recall that

[3]The S&P500 is a stock market index based on the market capitalizations of 500 leading companies publicly traded in the U.S. stock market, as determined by Standard & Poor's rating agency. The DJIA is based instead on the average of 30 publicly owned companies in the United States. Both are considered as the benchmark of the american stock market performance.

[4]Data provided by Bloomberg.

the kurtosis of a distribution $F(x)$ is defined as the fourth normalised cumulant [48] [5]

$$\kappa = \frac{\langle (x - \langle x \rangle)^4 \rangle}{(\langle x^2 \rangle - \langle x \rangle^2)^4}, \tag{2.3}$$

where $\langle \cdot \rangle$ stands for expected value defined as $\langle x \rangle = \int x f(x) dx$, with $f(\cdot)$ the probability density function associated with $F(\cdot)$. We can appreciate from Table 2.1 how κ tends to the benchmark value of 3 (expected for a Gaussian distribution [27]) as the time scale of returns is decreased, while it is larger than 3 for other smaller time scales. This feature reveals a universal property of stock returns: the distribution is far from being Gaussian in the high frequency regime (from the intra-day to the weekly returns), but still the Gaussian behaviour is recovered as one considers longer time scales. This means that the unconditional statistics of stock returns becomes more regular as one moves from short to long term investments. At the same time though, short term investments and high-frequency trading allow for bigger risks and potentially higher profits, a fact that reflects onto the tails of returns distribution being thicker. The S&P500 returns for the three different time scales $\tau = 1, 5, 25$ are plotted in Figures 2.2 and 2.3. One can see how the heterogeneity in the fluctuations size decreases along with the increase of the time scale.

2.2.3 Fat tails

The departure from normality is usually quantified by looking at the empirical cumulative distribution function (cdf) of stock returns, which for a density $f(x)$ is defined as [27]

$$F_<(x) = \mathbb{P}(X < x) = \int_{-\infty}^{x} f(u) du. \tag{2.4}$$

Stocks and indices are known to exhibit power law scaling in the cdf tail region [49, 50], that is their complementary cdf $F_>(x)$ displays the following behaviour

$$F_>(x) = 1 - F_<(x) = \mathbb{P}(X > x) \sim x^{-\alpha} \tag{2.5}$$

[5] Often one reports the excess kurtosis, defined as $\tilde{\kappa} = \kappa - 3$, $\kappa = 3$ being the kurtosis of a Guassian distribution.

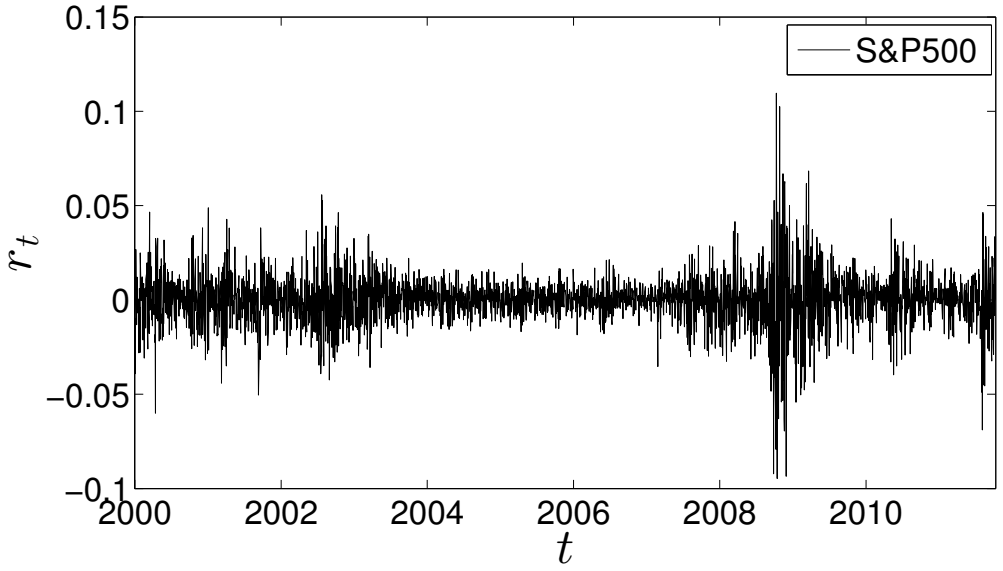

FIGURE 2.2: **Daily S&P 500 returns.**

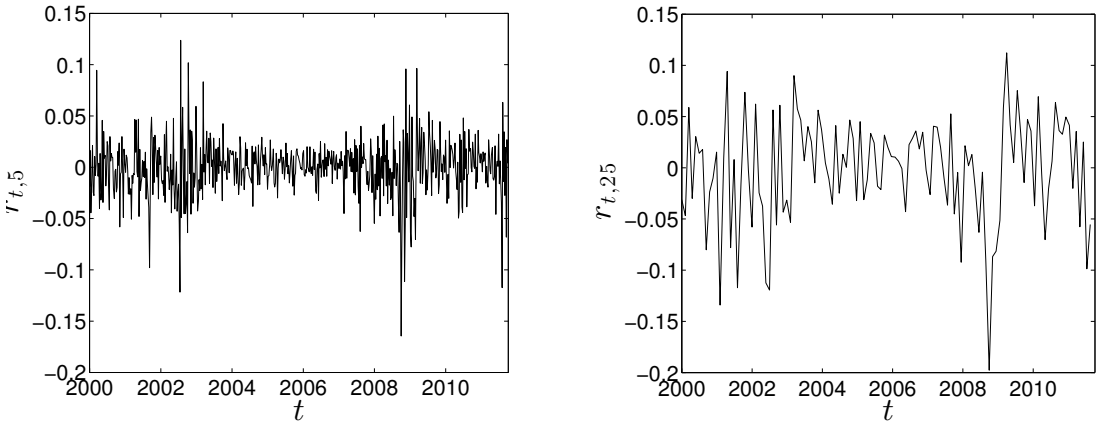

FIGURE 2.3: **S&P 500 returns at weekly (left) and monthly (right) scale.**

with a tail exponent $\alpha \in [1, 5]$ [27]. For a power-law cdf, events many standard deviations away from the mean are still likely to be observed and the tail exponent α quantifies the likelihood of such extreme fluctuations: the larger α, the closer the distribution is to a Gaussian[6], whereas $\alpha \in [1, 3]$ indicate stocks whose behaviour is far from being Gaussian and which show remarkably large fluctuations[7] [27]. In Figure 2.4 we report on the left examples of the power-law scaling observed on

[6]For a normally distributed random variable large deviations are very unlikely: as an example, a fluctuation of $10s$, with s the standard deviation, has a probability of 2×10^{-23}.

[7]Stock returns exhibiting $\alpha \leq 2$ have in principle infinite variance and are therefore the most unstable and risky ones.

single stocks and on the right the histogram of the observed tail exponents α on a dataset of 342 most capitalised stocks listed in NYSE in the period 1997-2012. The stocks in the left plot are: American International Group (AIG), First Horizon National Corporation (FHN), Abbot International (ABT) and General Electrics (GE). The left plot in Figure 2.4 is a so-called rank-frequency plot [51]. This is a very convenient and simple method to analyse the tail region of the distribution without any loss of information which would instead derive from gathering together data points with an artificial binning. In order to make this plot from a given set of observations $\{x_1, x_2, ..., x_T\}$, one first sorts the T observed values in ascending order and then plots them against the vector $[1, (T-1)/T, (T-2)/T, \ldots, 1/T]$. Indeed, we have that $\text{Rank}(x_i) = 1 - F_<(x_i)$. The tail exponent α is then estimated as average of several fits of the linear part of $F_>(x)$ in log-log scale, performed considering different values for the lower extremum x_{min} at which the fit is started: this also allows to associate an error to α. We mention that this method returns a tail exponent consistent with the one obtained by means of the Hill estimator [52, 53], another standard estimator for α. In the right plot of Figure 2.4 we also

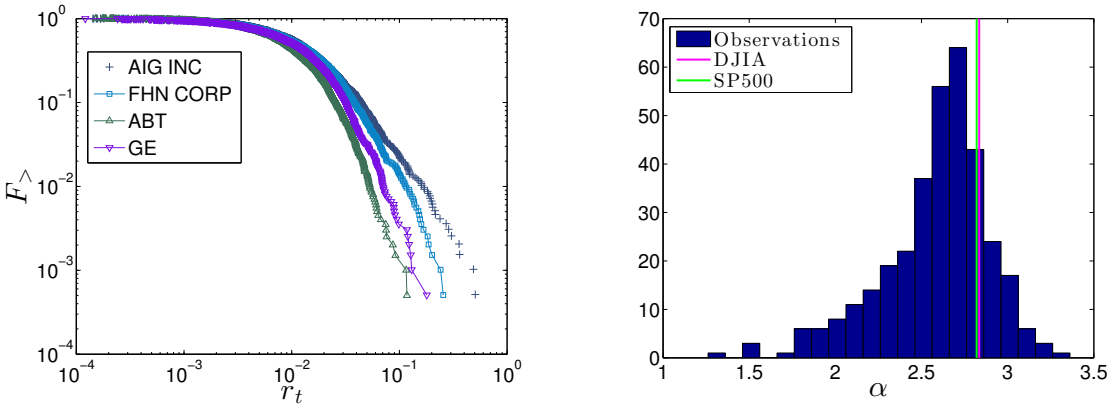

FIGURE 2.4: (Left) Rank Frequency plot in log-log scale for returns of four different stocks whose tickers are given in the legend. Mean and standard deviation of α for the four stocks obtained by varying x_{min} in the linear fit are: $\alpha = 1.85 \pm 0.02$ for AIG, $\alpha = 1.96 \pm 0.01$ for FHN, $\alpha = 2.76 \pm 0.02$ for GE and $\alpha = 3.29 \pm 0.02$ for ABT. (Right) Histograms of the mean values of α obtained on the entire data set 2. The vertical lines are the values of α obtained for the S&P 500 (green line) and DJIA (magenta line).

show the values of α retrieved for the two composite indices, which are pretty close to the mode of the histogram. Indices typically present values of α close to

TABLE 2.2: **We report in this table the firms with $\alpha \leq 2$ and corresponding business sectors given by Bloomberg.**

Stock	Sector
American International Group INC	Financial
Crown Holdings INC	Consumer Goods
Eastman Kodak Co	Consumer Goods
Freddie Mac	Financial
Fannie Mae	Financial
Fifth Third Bankcorp	Financial
Huntington Bancshares INC/OH	Financial
Lincoln National Corp	Financial
XL Capital LTD	Financial
General Growth Properties INC	Financial
Simon Property Group INC	Financial
Ambac Financial Froup INC	Financial
Hartford Financial Services Group INC	Financial
Avis Budget Group INC	Services
Wells Fargo & Co.	Financial
PG&E CORP	Utilities
SLM Corporation	Financial
YRC Worldwide INC	Services
MBIA INC	Financial
Magic Investment CP	Financial
HCP INC	Financial

3 and are seldom found to exhibit very large fluctuations, as any extreme event in the single stocks time series is averaged out. Hence, although being very useful to track the performance of the whole market, focusing on indices only may be misleading as much more extreme fluctuations and far-from-Gaussian behaviour are observed in single stocks. For this reason, we have focused the whole analysis reported in this thesis on stocks from different sectors, in order to spot remarkable statistical features at a micro level, rather than looking only at indices. We report in Table 2.2 the names of the firms exhibiting $\alpha \leq 2$ and their corresponding business sectors. One can see that there is a predominant component of firms from the Financial sectors, renown for including some of the more risky stocks in the market.

As already discussed in the previous paragraph, the kurtosis of stock returns increases along with the decrease of the time scale at which returns are considered. This feature is reflected in the tails of the corresponding distributions growing

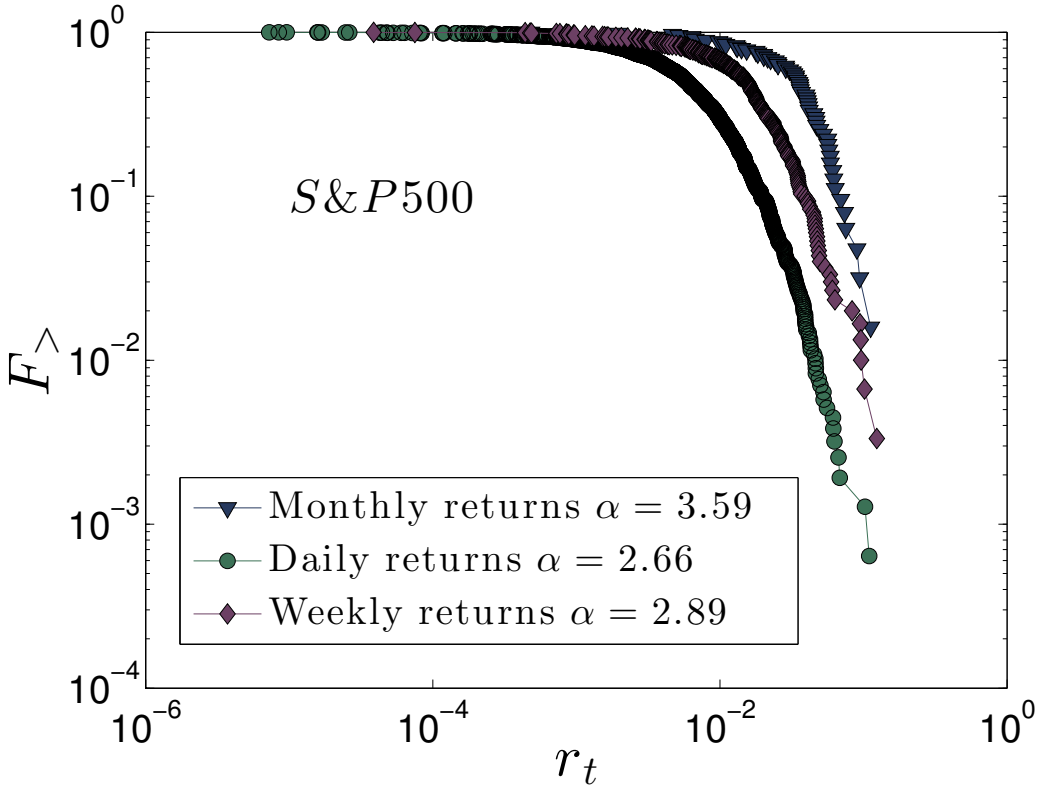

FIGURE 2.5: Rank-frequency plot in log-log scale for S&P500 index at daily (green circles), weekly (purple diamonds) and monthly (blue triangles) scale. The resemblance to the Gaussian distribution is enhanced considering a coarser scale. The errors on the α's reported in the legend are all equal to 0.03.

thicker as we decrease the time scale. In Figure 2.5 we show the complementary cdfs for S&P 500 returns at daily, weekly and monthly scale, which show very well how the distribution departs from the Gaussian the more we decrease the time scale.

2.2.4 Uncorrelated Returns and Volatility Clustering

Alongside the properties of the unconditional distribution explored in the previous paragraphs, stock returns exhibit remarkable temporal dependence [1, 2]. The time dependence reflects the memory of the market, by which large stock price swings at a certain time feed back on the stock time series and affect future price movements. This means that returns time series cannot be treated as an i.i.d. sample, but should be modelled in such a way that the observed temporal correlation be taken

into account. Looking again at the plot in Figure 2.2 we can see that, although returns do not show any particular trend, their variance strongly depends on time, alternating rather high levels in turbulent periods with lower levels in relatively calm periods. The credit crisis is particularly evident in the time period late 2007-late 2009, with S&P 500 returns taking on values never reached before. Computing the autocorrelation function of returns though, defined as [2]

$$C_r(\ell) = \langle r_t r_{t+\ell} \rangle - \langle r_t \rangle \langle r_{t+\ell} \rangle \sim \langle r_t r_{t+\ell} \rangle, \tag{2.6}$$

one finds that (for daily returns) this is zero after a very short lag. The same is found to hold for intra-day data [27]. We report in Figure 2.6 a plot of $C_r(\ell)$ for DJIA (left) and for several stocks (right). The observation of such a fast decay is usually explained in the framework of the efficient market hypothesis [54]: any time correlation in an observable variable (returns, in this case) would clearly give the opportunity to devise a strategy to make profits on future price movements. The absence of correlation in returns therefore explains how the market absorbs these strategies and does not allow for "easy lunch", apart from some very short lag correlations which correspond to the time needed for the market to process the new information.

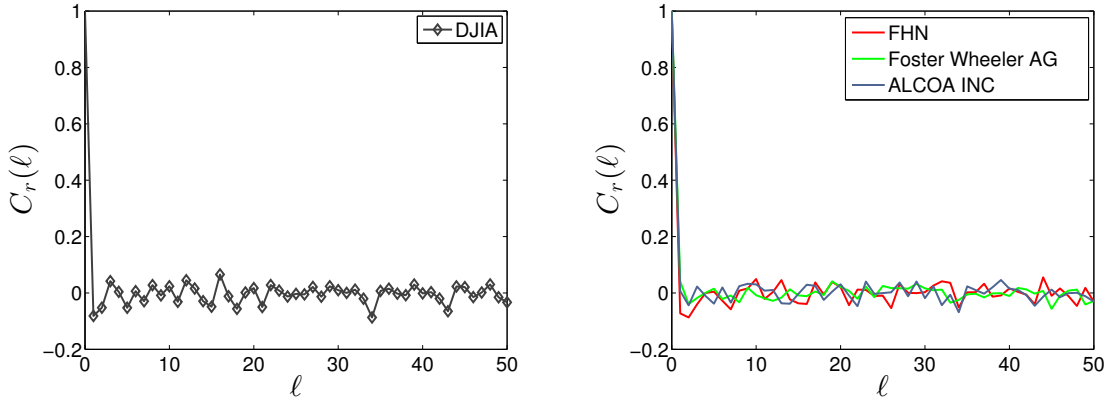

FIGURE 2.6: We plot return autocorrelation functions of DJIA daily returns in the period 1999-2011 (left) and daily returns of three stocks (right) in the period 1997-2012. Stocks are: First Horizon National Corporation, Foster Wheeler and ALCOA.

The presence of some form of persistence in returns though emerges as soon as

one starts to look for non-linear correlations. The most striking feature of many asset classes returns is in fact the so-called effect of *volatility clustering*, which reflects the heteroskedastic[8] nature of the variance. This effect was discovered for the cotton price and reported in [34] by Benoit Mandelbrot, who noted that *"large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes "*. As one can see again from Figure 2.2, large returns tend to lump together, alternating clusters of large fluctuations with clusters of small fluctuations. These patterns observed in returns reflect in a slow decay of several non-linear autocorrelation functions of returns $C_{f(r)}(\ell)$. The most notable examples of these non-linear functions are

$$\begin{aligned} C_{r^2}(\ell) =& \langle r_t^2 r_{t+\ell}^2 \rangle - \langle r_t^2 \rangle \langle r_{t+\ell}^2 \rangle, \\ C_{|r|}(\ell) =& \langle |r_t||r_{t+\ell}| \rangle - \langle |r_t| \rangle \langle |r_{t+\ell}| \rangle, \\ C_{\ln r}(\ell) =& \langle \ln |r_t| \ln |r_{t+\ell}| \rangle - \langle \ln |r_t| \rangle \langle \ln |r_{t+\ell}| \rangle. \end{aligned} \tag{2.7}$$

All these functions are found to be well approximated by power-laws of the form [26, 27]

$$C_{f(r)}(\ell) \sim A\, \ell^{-\beta} \tag{2.8}$$

with the parameter $\beta < 1$. The power-law decay (2.8) is shown for the three correlation functions in Equations (2.7) for Nike stock[9] in the period 1997-2012 in Figure 2.7. The decay exponent β is found to be $\beta = 0.1982$ ($C_r(\ell)$), $\beta = 0.1098$ ($C_{\ln|r|}(\ell)$) and $\beta = 0.0724$ ($C_{|r|}(\ell)$).

Similarly to what observed for the distribution tail exponents α, it is interesting to see how this slow decay in the autocorrelation functions $C_{f(r)}(\ell)$ depends on the time scale τ at which returns are evaluated. We report the ACF of square returns $C_{r^2}(\ell)$ for daily, weekly and monthly returns in Figure 2.8 for the DJIA index. We can see how the power-law decay is progressively lost as we move towards larger scales and the decay of $C_{r^2}(\ell)$ changes into exponential. This means

[8] The term heteroskedasticity denotes a property of a stochastic process whose variance is not constant in time.

[9] Data from Bloomberg.

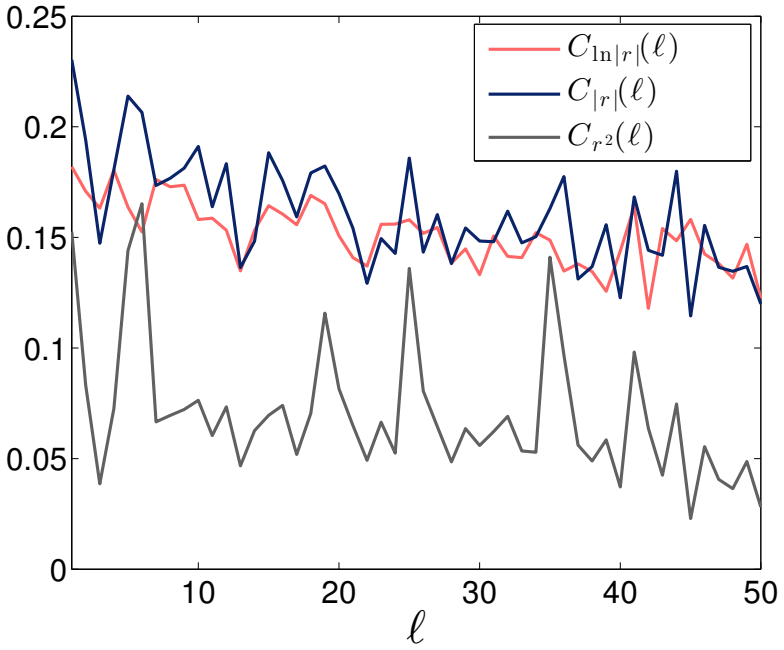

FIGURE 2.7: **Non-linear correlation functions of daily returns for Nike stock in the period 1997-2012.** We plot the autocorrelation functions of three non-linear functions of returns given in Equations 2.7.

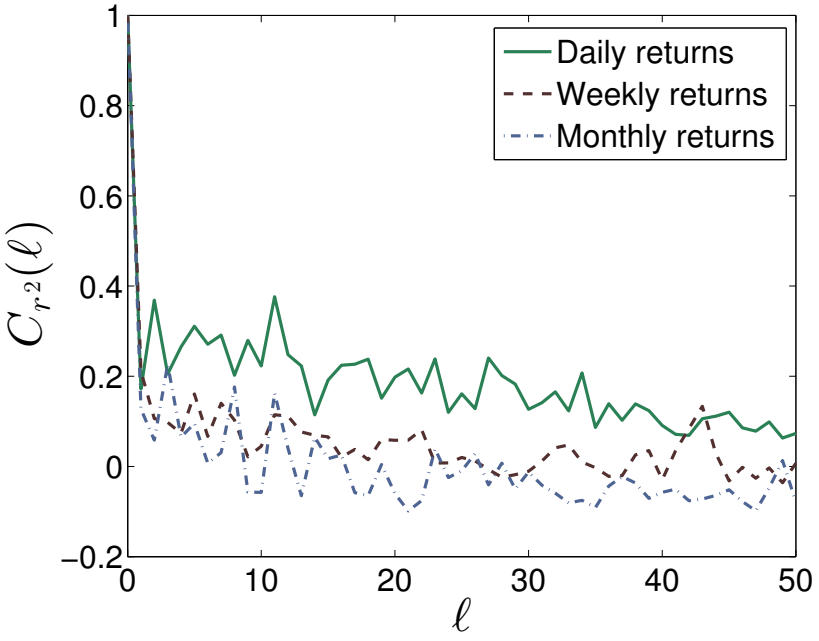

FIGURE 2.8: **Autocorrelation function of square returns for DJIA index at different time scales.** The long-range dependence is progressively lost as the return time scale is increased.

that the volatility of low-frequency returns is less persistent than that of high frequency returns and therefore low-frequency returns are better approximated by

an i.i.d. process. These observations, together with those already pointed out on the unconditional distributions of returns, allow to give the following summarising remark about asset returns: the high frequency regime (from intra-day to few days) displays highly non-linear properties characterised by non-Gaussian distributions (excess kurtosis $\tilde{\kappa} > 0$) and long-range correlated volatility. Conversely, low frequency returns (from weekly returns to quarterly) show properties much more consistent to a standard random walk, where returns can be considered to be independent and their distribution to be Gaussian. From this very general observation one can quickly grasp why modern finance is very much concerned with high-frequency trading: the statistics of returns at high frequency offers a much richer field for seeking extreme risk and (potentially) huge gains than the flat behaviour of low-frequency finance, where large fluctuations and temporal correlations are all but absent. The downside of this is that these hidden gains come at the very high risk embedded in these highly non-linear systems, where price drops are also much more likely than in a linear Gaussian framework.
Therefore, the higher the frequency at which returns are computed, the more complex the behaviour of the resulting time series. Both the tail exponent and the various decay exponents for different non-linear autocorrelation functions of returns represent a direct measure of this intrinsic complexity of stock returns. As it turns out however, there is another aspect, more general, including both the two stylised facts analysed so far, which can allow to track the complexity of stock returns with a single number: this aspect is called *multifractality* and we investigate it in the next section.

2.3 Multifractality of Stock Returns

Multifractality (or multi-scaling) of a stochastic process is a concept related to the way this process scales across different time scales [32]. It reflects the non-linear relation between the properties of the distribution of the process increments at different time scales. It is associated with the complexity of the process, and any proxy of multifractality should be an increasing function of the complexity

and vice-versa. The concept of multi-scaling has its origin in physics, but has been investigated and applied more and more often in finance starting from the pioneering studies of Mandelbrot [55–57]. The last twenty years in particular have seen the appearance of many studies on multifractal properties of asset returns [8, 58–66] and many models have been proposed in order to capture the regularities exhibited by the available data [20, 28, 67–69].

A stochastic process S_t with stationary increments is called multifractal if the following scaling law is observed

$$\mathbb{E}(|S_{t+\tau} - S_t|^q) \sim c_q \tau^{\zeta_q}, \tag{2.9}$$

where c_q is a constant which depends on q and ζ_q is a non-linear function of q, called the *scaling function*. Defining $\delta_\tau S_t = S_{t+\tau} - S_t$ one sees that Equation (2.9) is a scaling relation for the q-moments of the increments of the process S_t. In fact

$$M_q(t,\tau) = \mathbb{E}(|S_{t+\tau} - S_t|^q) = \int_{\mathbb{R}} |\delta_\tau S_t|^q dP_S, \tag{2.10}$$

where P_S is the probability distribution associated with the process increments. Note that the presence of the absolute value ensures that even for $q = 1$ the variable $|\delta_\tau S_t|$ has non-trivial temporal correlations, as seen in the previous section. Equation (2.9) thus relates the scaling of the moments of the S_t-distribution across different time scales τ and the inhomogeneous scaling of moments of different orders implies that the distribution of the process increments critically depends on the time scale of the increment. What distinguishes multifractal processes from simply fractal ones is the scaling being *non-linear* and the scaling function ζ_q governs this non-linearity.

The scaling in Equation (2.9) has been shown to hold robustly in financial markets across different asset classes [28, 58, 70], from composite indices and stocks to currency and foreign exchange [13, 38, 39, 71]. The multifractal behaviour is typically visualised by looking at the scaling of empirical moments, computed as

$$M_q^*(t,\tau) = \frac{1}{T-\tau+1} \sum_{t=0}^{T-\tau} |r_{t,\tau}|^q \sim \tau^{\zeta_q^*}, \tag{2.11}$$

where T is the time-series length and ζ_q^* is the measured exponent of the scaling relation. It is well known [72, 73] that the empirically estimated ζ_q^* is significant only for small values of q and one therefore needs to be careful in interpreting the scaling beyond a certain q. Usually the estimation is considered unreliable above $q = 4$. Moreover, for multifractal processes it is known that the scaling in Equation (2.9) holds for small τ only, more precisely for $\tau/T \ll 1$, where T is some larger scale called the *integral scale* [74]. We show in Figure 2.9 two examples of the scaling predicted by Equation (2.9). The non linearity of the scaling with q is evident from the different slopes of the lines corresponding to different values of q. Note that for $\tau \to \infty$ all lines converge to a single value as the scaling functions

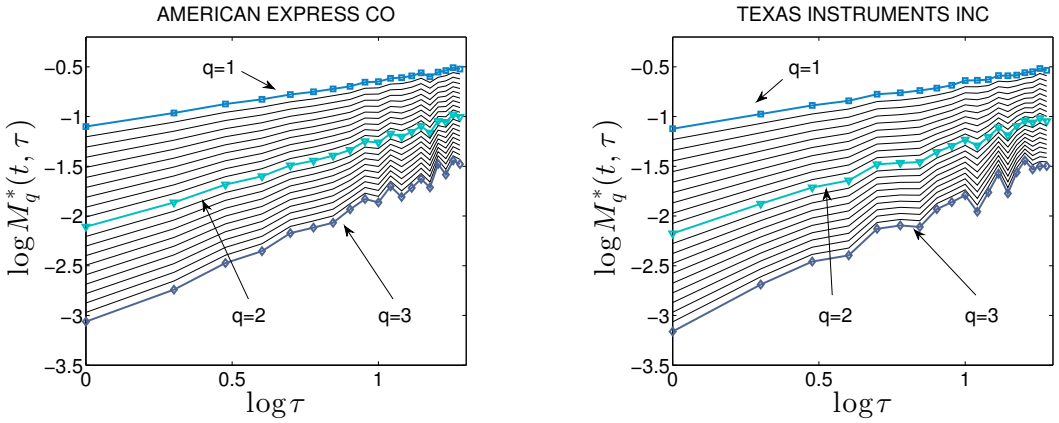

FIGURE 2.9: We plot the logarithm of the empirical moments against the logarithm of the time scale τ. (Left) American Express Co daily returns. (Right) Texas Instruments INC daily returns. The different slopes of the lines corresponding to different values of q reveal the multifractal properties of the data.

for different q collapse to one value.

In the Econophycis literature, the two most widely used methods to directly quantify the multifractal properties of a time series are the following: the Multifractal Detrended Fluctuation Analysis (MF-DFA) [15] and the Generalized Hurst Exponent (GHE) [38]. Both methods aim at estimating the scaling function ζ_q by means of a scaling exponent, which is a q-dependent generalisation of the Hurst exponent H. Barunik and Kristoufek [75] show that the GHE method outperform the MF-DFA in returning an estimate of the scaling exponent with lower variance and bias regardless of the presence of heavy tails in the data. For this reason

we have chosen to use this method in order to give a quantitative description of multifractality. We describe the GHE method in the next section.

2.3.1 Generalized Hurst Exponent method

The GHE method is a tool to directly study the multifractal properties of financial time series by looking at the scaling of the q-moments [76]. The GHE $H(q)$ is defined via the scaling of Equation (2.9) as

$$\zeta_q = qH(q). \tag{2.12}$$

We remark that this definition holds conditionally on the scaling being actually observed. If the empirical moments scale according to Equation (2.11), then the GHE is computed by averaging several linear fits of the relation

$$\log M_q^*(t,\tau) = qH(q)\log\tau + \log c_q^* \tag{2.13}$$

where each fit is performed with a different $\tau \in [1, \tau_{\max}]$. For daily data, τ_{max} is typically varied between 5 and 19 [13].

Equation (2.12) allows to distinguish between two types of processes: (1) processes where $H(q) = H$ is constant independent of q and (2) processes where $H(q)$ depends on q. The first class is that of uni-fractal (or simply fractal) or uni-scaling processes, for which all the lines in Figure 2.9 have the same slope and moments of different order scales with the same exponent H. The second class is that of multifractal (or multi-scaling) processes and $H(q)$ quantifies the scaling of each different order of moments. The scaling function computed through the GHE can be plotted against q and, for financial data, typically reveals concave shape like those plotted in Figure 2.10. In the same figure we also show the comparison with a uni-scaling process having a constant H and thus a linear scaling function ζ_q.

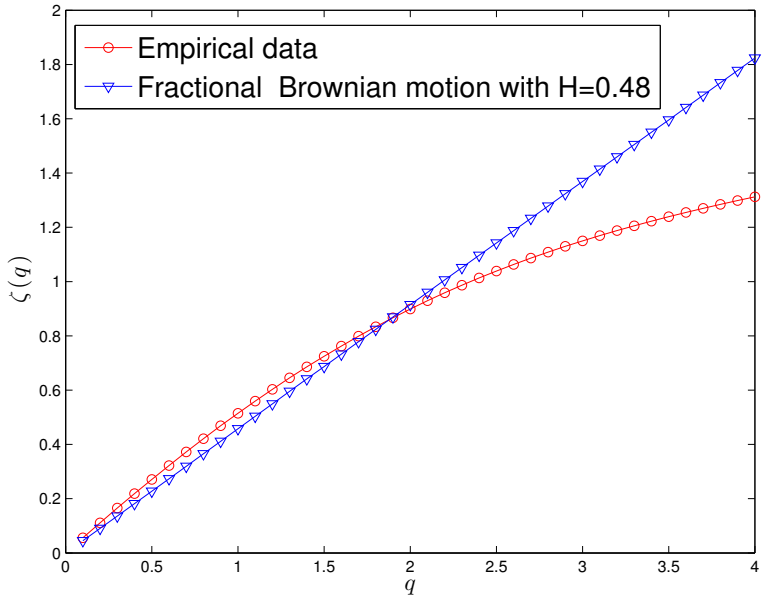

FIGURE 2.10: We plot in red circles the scaling function computed via the GHE method for America Express stock and in blue triangles the scaling function for a Fractional Brownian Motion with Hurst index $H = 0.48$, whose linearity reveals the uniscaling nature of the process.

2.3.2 GHE and Market Stage of Development

The economic relevance of the GHE has been first pointed out in the publications [38, 39]. The authors have shown, indeed, that there is a remarkable correlation between the value of the GHE and the stage of development of a market. Well developed markets, like the American, the Japanese and major European ones, systematically exhibit GHE $H(q) < 0.5$ for $q = 1, 2$, whereas emerging markets tend to show $H(q) > 0.5$ for $q = 1, 2$. This results is found for Indices as well as single stocks and currencies, proving to be very robust across very different assets [13]. $H(q) < 0.5$ are associated with persistent processes, while $H(q) > 0.5$ with anti-persistent ones. The reference value of $H(q) = H = 0.5$ corresponds to a Brownian motion with no multi scaling properties.

The observation of such a neat differentiation across the scaling exponents has the following interpretation, also proposed in [2]: the development of a market has its ground in the heterogeneity of the frequency at which different agents operate in it. Developed (and efficient) markets are characterised by a wide range of frequency

of market operations and investments, which enrich the dynamics endogenously and make the resulting market more complex and structured. Emerging markets on the other hand are often characterised by the absence of some types of agents, which renders their structure inherently different. The relevance of scaling analysis thus lies in its ability to characterise the volatility at different time scales, which allows to parsimoniously assess the impact of the heterogeneous agents on the prices. All the information is enclosed in one number, the GHE, which can be extremely useful as a benchmark for institutional investors looking for the best investment strategies.

2.4 The Source of Multifractality

In this section we show an original contribution to the unveiling of different contributions to multifractality. We analyse the set of daily prices of the 342 most capitalised stocks listed in NYSE in the period 2/01/1997 to 31/12/2012. Data have been provided by Bloomberg.

As we have seen in Section 2.3, multifractality depends on the way the distribution of returns behaves at different time scales and it is therefore influenced by both the temporal structure and the unconditional distribution of returns. The aspect which is worth to point out is that these two different empirical features are fundamentally associated to two different types of scaling: the first one concerns the scaling of any volatility measure (square or absolute returns for example) with the returns time scale; the second quantifies the scaling of the returns distribution with the size of the fluctuation, for a fixed time scale τ. Multifractality thus merges two different scaling behaviours and allows to quantify the complexity of the time series by means of a single variable, be it the scaling function or the GHE. A natural question which arises is then what is the relative weight of the different component in the overall multifractal properties of stock returns. This problem was first tackled in [15], through the MF-DFA on synthetically generated data with both short and long-range correlations and fat-tails. The multifractal spectra for

these synthetic series were compared to those obtained on the corresponding shuffled time series. If multifractality is only due to temporal correlation, then the shuffling procedure should destroy any signature of it. The authors found that for synthetic series showing long-range correlations *and* broad probability distributions multifractality could not be completely destroyed by shuffling the time series, whereas for long-range correlated process without fat tails any signature of multifractality would completely disappear after the shuffling.

More recently Zhou [16] has provided a very detailed account of the different components of multifractality of the DJIA index daily returns by comparing its properties with differently generated synthetic series, twisting and tweaking their long-memory and fat tails properties. Overall, he found that the major impact on the singularity spectrum of the MF-DFA (a proxy of multifractality) is caused by the returns distribution having fat tails, while the temporal structure played a minor role, confirming previous results already shown in [40, 41].

Further investigations about the source of multifractality have been reported in [17], this time through the GHE method. In this paper the authors have observed the puzzling effect of having multifractality in the analysed series *increased* after data shuffling. They have interpreted this apparently queer behaviour as arising from the appearance of very short ranged correlations in the shuffled series which contribute to multifractality more than long-range correlations do.

2.4.1 Separating the two contributions

We inspect the bizarre effect reported in [17] a bit more in detail. Starting from the scaling exponents $H(q_1)$ and $H(q_2)$ we can define a proxy of multifractality as

$$\Delta H(q_1, q_2) = H(q_1) - H(q_2), \qquad q_1 \neq q_2. \tag{2.14}$$

For practical purposes and to have robust statistical estimations of the scaling exponents we stick to $q_1 = 1$ and $q_2 = 2$ [10]. Since, as we have discussed, multifractality is supposed to have two contributions, assuming these two contributions are completely separable, for a fixed pair of (q_1, q_2) Equation (2.14) can be decomposed further as

$$\Delta H(1,2) = \Delta H = \Delta H_\alpha + \Delta H_{LR}, \tag{2.15}$$

where ΔH_α is the contribution coming from the distribution and ΔH_{LR} is that from long-range dependence. After the time series have been reshuffled, any contribution coming from temporal dependence in the data is lost and one expects to find only properties of the unconditional distribution, that is

$$\Delta H_S \simeq \Delta H_\alpha \tag{2.16}$$

where with ΔH_S we label multifractality of shuffled data. We can then define the ratio

$$\delta = \frac{\Delta H_S}{\Delta H} \simeq \frac{\Delta H_\alpha}{\Delta H_\alpha + \Delta H_{LR}}. \tag{2.17}$$

Now, if the two contributions in Equation (2.15) are both positive, one would expect to have $\delta < 1$. This is actually not the case, as we show in Figure 2.11, where we plot δ for all the stocks in our dataset. In agreement with [17] we actually observe $\delta > 1$ in most of the cases, with $\langle \delta \rangle \sim 2.4$, which means that multifractality is higher for shuffled data than for the original time series. Further to the interpretation given in [17] this phenomenon has two other possible interpretations: (i) the contribution ΔH_{LR} is negative and removing it increases the overall degree of multifractality (ii) the decomposition in Equation (2.15) does not hold and the interplay between fat tails and long-range dependence cannot be split into two additive contributions to multifractality.

The first interpretation entails $H_{LR}(2) > H_{LR}(1)$, that is the absolute variance scales with an exponent larger than the absolute mean. On real market data, one observes most of the time $H(1) > H(2)$ and negative values of $\Delta H(1,2)$ are

[10] $q = 3$ also provides good estimation of the GHE but for heavy-tailed data is sometimes less reliable.

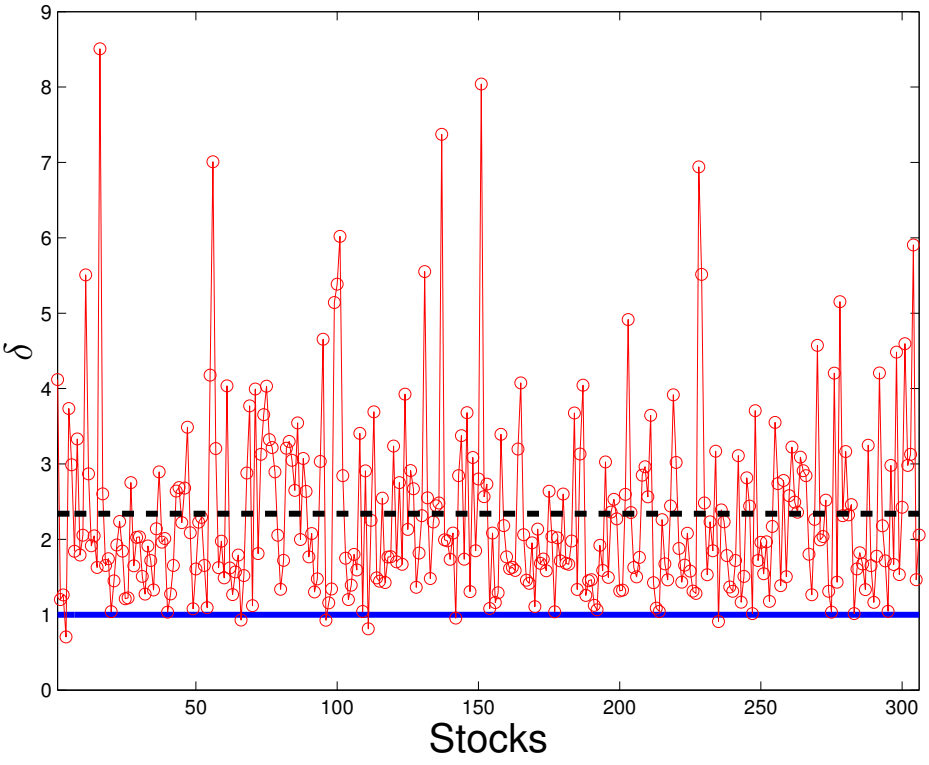

FIGURE 2.11: We plot the $\boldsymbol{\delta}$ ratio for all stocks in the dataset. The red circles are values for the single stocks while the black dashed line is the average $\langle\delta\rangle$. The blue solid line at $\delta = 1$ is the value expected for a scenario where data shuffling does not affect multifractality.

usually not statistically significant. On our dataset for example, only five stocks show negative $\Delta H(1,2)$ and none of these is found to be statistically different from $\Delta H(1,2) = 0$[11]. There is still no evidence though to exclude the possibility of this being actually the cause for the observed increase in $\Delta H(1,2)$.

Interestingly, as the plots in Figure 2.12 enlighten, one observes $\delta > 1$ also on synthetic multifractal time series simulated via Multifractal Random Walk (see Chapter 3 for details) and Markov-Switching Multifractal Model [12] [28] (see Appendix A for details of the model), whereas $\delta \sim 1$ is recovered on time series with

[11] The significance is tested against the hypothesis of unifractal processes, where deviations from $\Delta H(1,2) = 0$ are still possible due to statistical fluctuations. We simulated 1000 realisations of Fractional Brownian motions - a uniscaling process - and computed $\Delta H(1,2)$ each time. From the distribution of these values we found that all the five negative values are within one standard deviation from the mean $\Delta H(1,2) = 0$.

[12] The simulation reported in this paper has been run with parameters set as $K = 20, b = 2, m_0 = 1.5, \gamma_k = 0.5$ and $\sigma = 1$. For details about the meaning of each parameter see Appendix A.

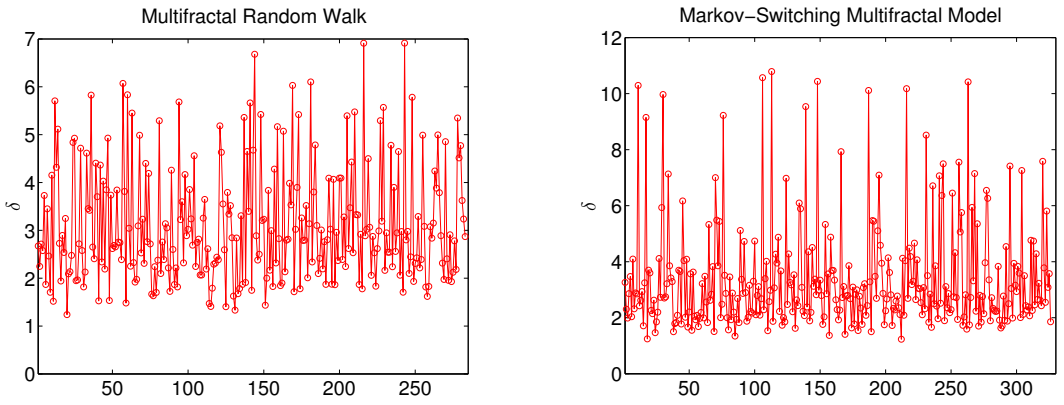

FIGURE 2.12: We plot the $\boldsymbol{\delta}$ ratio for (left) MRW time series simulated with $\lambda = 0.2, \sigma = 1$ and $T = 600$; (right) MSMM time series simulated with parameters $K = 20, b = 2, m_0 = 1.5, \gamma_k = 0.5$ and $\sigma = 1$. We have removed all time series whose measured multifractality cannot be considered statistically different from 0.

fat-tails but no long-range dependence. In Figure 2.13 we plot δ computed on Student-t distributed time series: the plot confirms that the effect of data-shuffling does not affect the multifractal properties, since the only contribution to the process's multifractality comes from the broad probability distribution. Moreover, as one can see from the range of variation of δ for Student-t variables compared to the ranges for empirical data and for simulated multifractal series, the hypothesis $\delta = 1$ is clearly met even without a rigorous statistical test.

The second interpretation is the one that sounds more reasonable, i.e. that long-range correlations and fat tails are two inherently entangled phenomena and their respective contributions to the multifractal proxy $\Delta H(1, 2)$ are non-additive.

The observations reported here however support the fact that the first responsible for the observed multifractality in financial markets be the unconditional distribution of returns and that, as also pointed out in [68, 77], stochastic volatility alone can reproduce apparent multifractal processes, even without calling in a rigorous cascade picture.

In the next section we are going to study the dynamical behaviour of scaling exponents and multifractality.

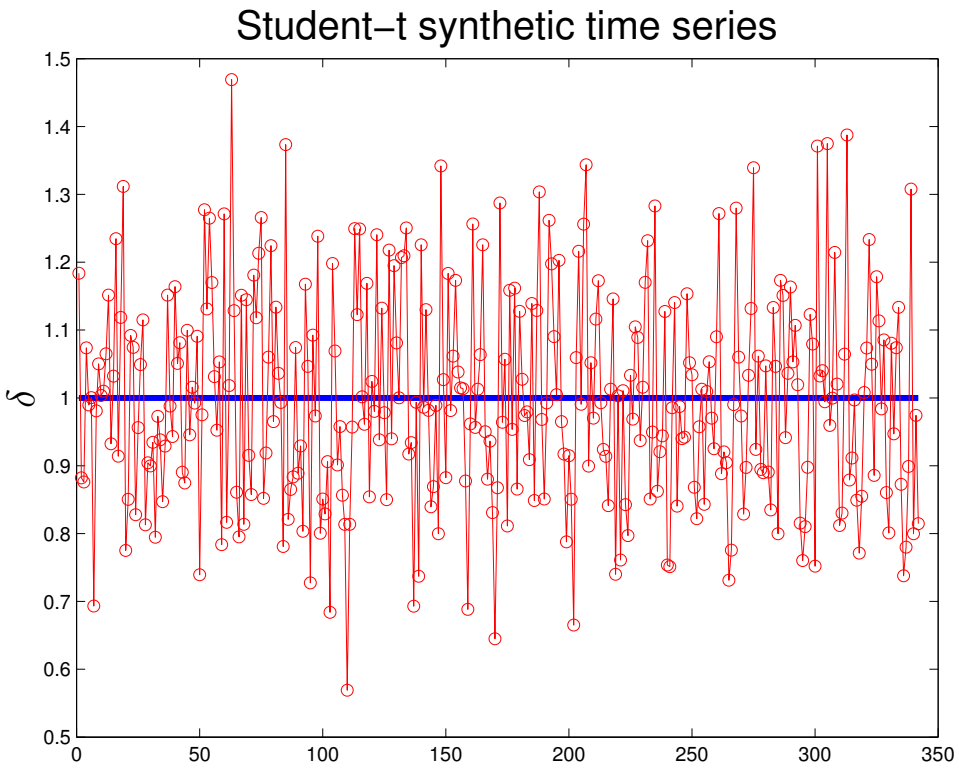

FIGURE 2.13: $\boldsymbol{\delta}$ ratio plotted for simulated Student-t time series with $\nu = 3.5$. As one can see, shuffling the data does not affect multifractality. The blue solid line is the average $\langle \delta \rangle$.

2.5 Dynamical weighted Generalized Hurst Exponent

In this section we introduce the weighted GHE (wGHE) and study its properties dynamically in time. We have concentrated on firms bailed-out in 2008 as a consequence of the credit crisis, but we have also considered other firms from different market sectors.

The dynamics of the GHE has been scarcely investigated in the literature, with some notable examples [78–82]. From a technical perspective a dynamical study of multi-scaling properties is very challenging because one must analyse a long enough time period to obtain a reliable statistics, but simultaneously the time-period must be short enough to catch changes occurring at given times. For this reason we have introduced exponential smoothing on a moving time-window and

studied the best combination of window size and exponential weights that satisfy both these requirements.

We analysed a dataset of 395 most capitalised companies listed in NYSE, considering daily prices in the period 1/01/1996 through 30/04/2009. Data were provided by Thomson Reuters.

2.5.1 Weighted Generalized Hurst Exponent

To take into account the fact that the recent past is more important than the remote past we can assume that the informational relevance of observations decays exponentially. This 'exponential smoothing' is attained by defining weights as

$$w_s = w_0 \exp\left(-\frac{s}{\theta}\right), \ \forall\, s \in \{0, 1, 2, \ldots, \Delta t - 1\} \tag{2.18}$$

where θ is the weights' characteristic time and its inverse is the exponential decay factor $\alpha = \frac{1}{\theta}$. The parameter w_0 is given by [83]

$$w_0(\alpha) = \frac{1 - e^{-\alpha}}{1 - e^{-\alpha \Delta t}}. \tag{2.19}$$

The weighted average over the time-window $[t - \Delta t + 1, t]$ for a general quantity $f(x_t)$ is thus

$$\langle f \rangle_w(t) = \sum_{s=0}^{\Delta t - 1} w_s \, f(x_{t-s}) \tag{2.20}$$

and the weighted GHE (wGHE) is therefore obtained by substituting the normal averages in Equation 2.13 with weighted averages:

$$M_q(t, \tau) = \langle |r_{t,\tau}|^q \rangle_w \sim \sum_{s=0}^{\Delta t - 1} w_s |r_{t,\tau}|^q. \tag{2.21}$$

From the scaling law in Equation 2.11 this leads to the linear relation

$$\log M_q(t, \tau) \sim q H^w(q) \log \tau. \tag{2.22}$$

from which the wGHE can be computed.

2.5.2 Results

In analogy with [38, 39] we have estimated the $H^w(q)$ as an average of several linear fits of Equation 2.22 with $\tau \in [1, \tau_{max}]$ and varying τ_{max} between 5 and 19. As proxy of the statistical uncertainty of the scaling law we have computed the standard deviation of the $H^w(q)$ over this range of τ_{max}. To track the evolution of the stage of development of a certain company, we have studied the dynamics in time of the wGHE on overlapping time-windows with a constant 50 days shift between any two successive windows.

First of all, to fully capture the advantages of the weighted average method, a choice of the parameters θ and Δt, namely the characteristic time and the width of the time-window, has to be made. In particular the time-window Δt must be large enough to provide good statistical significance but it should not be too large in order to retain sensitivity to changes in the scaling properties occurring over time. In order to satisfy both these requirements we took a rather long time-window Δt combined with a not too short short characteristic time θ. For instance, in Figure 2.14 we show how the manipulation of the parameter Δt affects the dynamics of the Hurst exponent of the company American International Group (AIG). As it can be appreciated from the figure, which shows plots for AIG with $\Delta t = 200$ days (left panel) and $\Delta t = 400$ days (right panel) respectively while keeping $\theta = 300$ days, the shape of the outline shrinks and gets neater as the time-window is increased. The left panel of Figure 2.14 shows more noisy dynamics when Δt is smaller. Conversely, in the right panel we can appreciate that a slimmer outline is achieved by increasing the statistics, but duly weighting it. We plot in Figure 2.15 the example of American International Group for the choice $\Delta t = 850$ and $\theta = 750$. In this plot the thick lines are the average $H^w(1)$ and the bands are given by the standard deviations over τ_{max} between 5 and 19 days [38, 39]. This choice of the parameters is probably the best as it allows to have a sufficiently large statistics, while still allowing to perform the analysis on the moving windows, but at the same time the events are weighted such that not all information present in the time series is given the same importance but at the same time no relevant information from the past is lost.

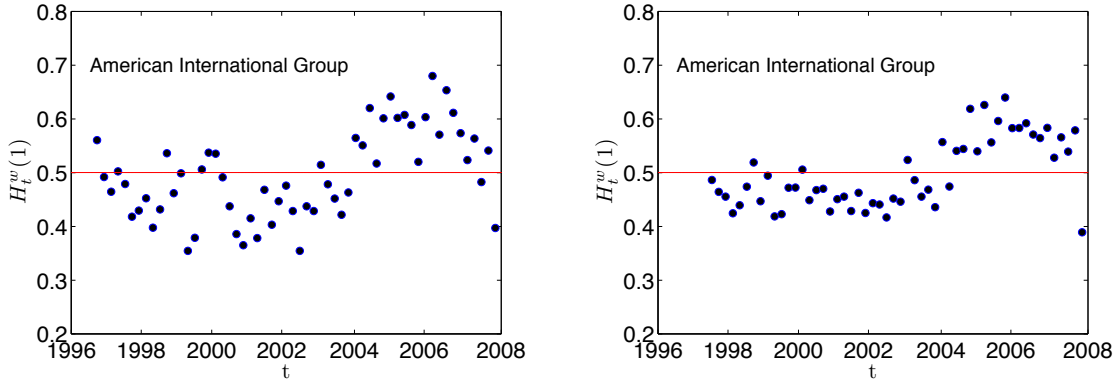

FIGURE 2.14: Weighted Generalized Hurst exponent $H^w(q=1)$ as a function of time for American International Group (AIG). Left panel: $\Delta t = 200$ days time-window. Right panel: $\Delta t = 400$ days time-window. The characteristic time is kept constant at $\theta = 300$ days in both plots. The points are reported in correspondence of the end of the time-window.

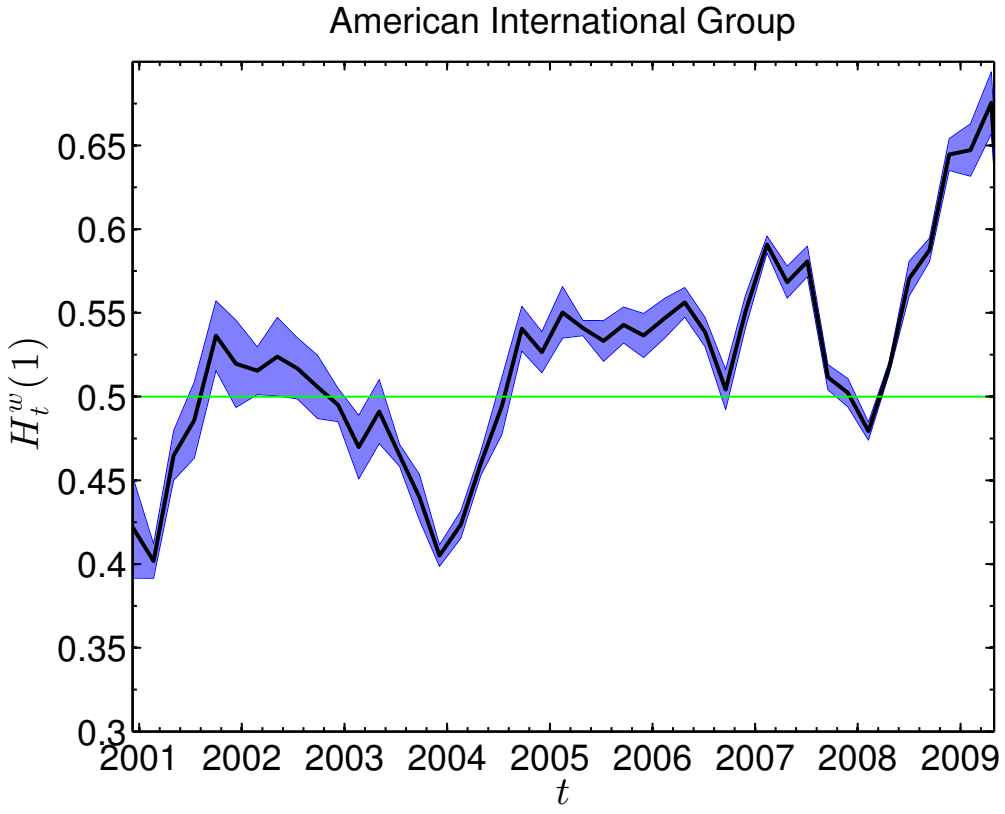

FIGURE 2.15: Weighted Generalized Hurst exponent as a function of time for American International Group (AIG). The overlapping time-windows are $\Delta t = 850$ days, with $\theta = 750$ days. The values are plotted in correspondence of the end of the time-window (solid black line). The shaded areas around the tick-line plot represent the sizes of the standard deviations.

Once the choice of the parameters is made, one can see that $H^w(1)$ for AIG is increasing in the period under study, with a transition from values < 0.5 to values > 0.5. For the bailed-out companies it would also be interesting to look at $H^w(2)$, which, as we said, is associated to the scaling of the auto-correlation function of the time series. However, in spite of the behavior being very similar to that observed for $H^w(1)$, for $\alpha < 2$ (which is the case for these companies) the second moment is not defined and thus it's difficult to interpret the real meaning of $H^w(2)$.

In Figure 2.16 the dynamics in time of $H^w(1)$ for the companies Freddie Mac and Fannie Mac is reported. These are public government sponsored enterprises which in September 2008 had to be put into conservatorship by the U.S. Treasury; namely the huge debts of these companies were purchased by the U.S. government. After playing a central role in the market during the mortgages's boost both firms defaulted. Their fate is pretty well pictured by the dynamical wGHE. Indeed,

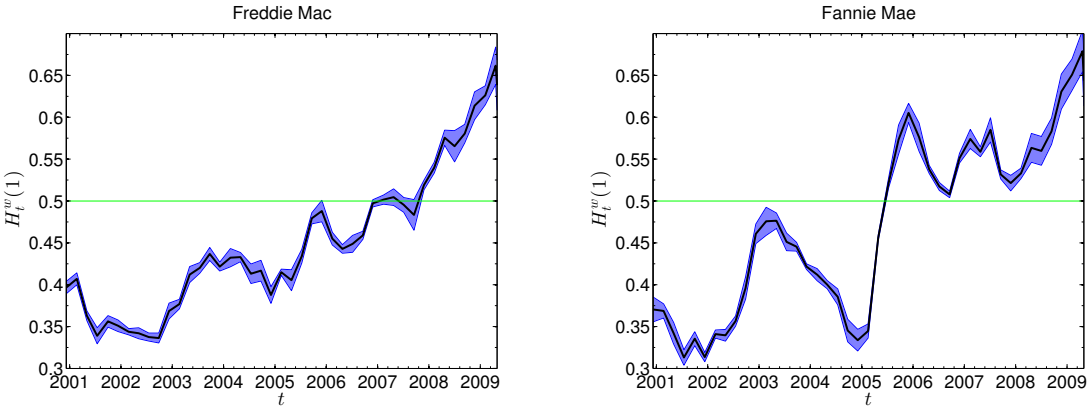

FIGURE 2.16: Weighted Generalized Hurst exponent $H^w(q = 1)$ as a function of time. Left panel: Freddie Mac. Right panel: Fannie Mae. The increasing trend over the whole period highlights a transition from values of $H^w(1) < 0.5$ to values of $H^w(1) > 0.5$. This suggests a progressive change in the stability of the companies under study.

there is a clearly visible trend in these plots showing how the value of $H^w(1)$ for these companies has been - on average - increasing since 1996 until 2009. According to [13, 39] these trends might suggest a transition from a stable stage of the companies to an unstable one.

Other bailed-out companies which show the same trend are shown in Figure 2.17. Again the trend is increasing and crossing over the value of 0.5 towards the end of the time-period when the crisis fully unfolded.

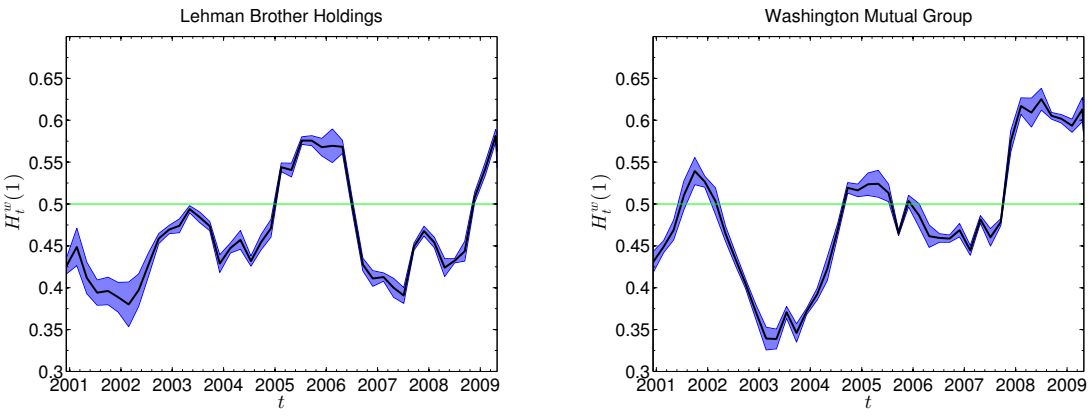

FIGURE 2.17: Weighed Generalized Hurst exponent $H^w(q = 1)$ as a function of time. Left panel: Lehman Brothers Holding, the bank who defaulted on 15th September 2008, probably the key event of the credit crisis. Right panel: Washington Mutual. The increasing trend over the whole period highlights a transition from values of $H^w(1) < 0.5$ to values of $H^w(1) > 0.5$. This suggests a progressive change in the stability of the companies under study.

We have compared these results with those obtained by looking at other companies either from the financial sector or belonging to other market sectors to test the significance of these results. For example, in the Basic Materials sector, we found many companies whose dynamical wGHE decreases in time, thus exhibiting an opposite behavior to that shown by the bailed-out companies in the financial sector. An example is reported in Figure 2.18 where the dynamical wGHE's for two companies belonging to the sector of Basic Materials are shown. We notice a very definite overall decreasing trend, as if the companies securities gained persistence in going through the period of crisis. This is in agreement with what has been considered as the boost of the commodities market during the crisis [84, 85], where investors were turning away from the financial sector.

There are other sectors that have revealed instead no particular trend in the dynamical wGHE. We stress that even in the Financial sector itself, the increasing trend found for the bailed-out companies is not common to others; for instance, many companies, like American Express Co and Morgan Stanley show stable behaviors, with wGHE values steadily fluctuating about 0.5. We will see in the next section that the sectors exhibiting a definite trend in the dynamical wGHE are also those showing extreme values in the tail exponents of their distributions of returns. Although the increase or decrease of the wGHE is not simply related with

the return statistics only, both behaviors are associated with the fluctuations of the log-returns distributions.

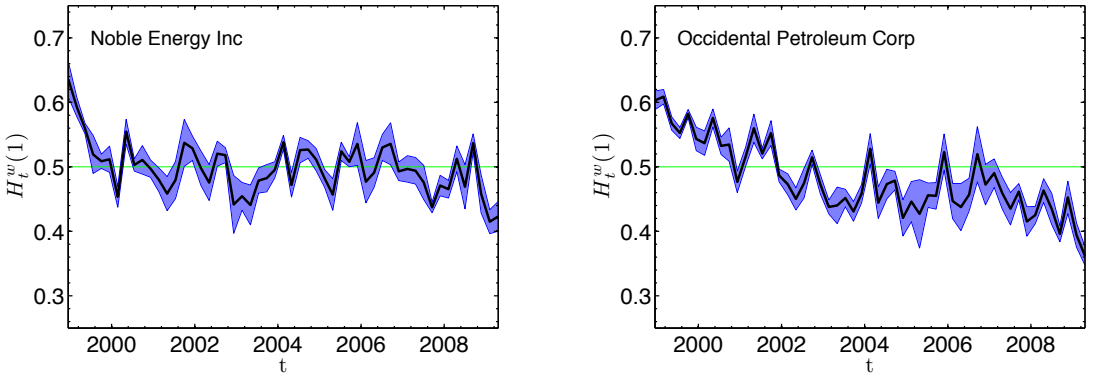

FIGURE 2.18: Weighed Generalized Hurst exponent $H^w(q=1)$ as a function of time for: Left panel - Noble Energy Inc.; Right panel - Occidental Petroleum. The time-window is taken to be $\Delta t = 850$ days and $\theta = 750$ days.

The entire analysis reported above has been also performed on $H_t^w(2)$ revealing very similar behaviour, although on average the scaling exponent for $q = 2$ is found to vary less than that corresponding to $q = 1$.

The behaviour observed for the bailed-out firms is interpreted as a progressive loss of stability of these firms as investors gradually turned away from risky stocks. The trends spotted hence suggest the possibility of using the wGHE to track the level of stability of a firm in time.

2.5.3 Fat tails and extreme events

Further to the wGHE we have also looked at the tail exponents of the bailed-out firms. All corresponding time series show very large fluctuations, not accountable for by a Gaussian statistics. The tail exponents of the complementary cdfs are $\alpha \sim 1.7 \pm 0.03$ for AIG and Lehman Brothers and $\alpha \sim 1.6 \pm 0.03$ for Freddie Mac and Fannie Mae. We found extremely large kurtosis for the distributions of these stock returns (139 for AIG and 761 for LBH), confirming the fact that extremely large fluctuations are observed for these firms.

Excluding the recent unstable period from the same dataset though, i.e. taking off

the years 2008-2009, a slightly different picture emerges with the scaling exponents exhibiting larger values and the frequency of very large fluctuations becoming an order of magnitude smaller. Figure 2.19 shows the exponents for all the firms, computed both over the entire period and over the period excluding the crisis. We plot in different colours the exponents corresponding to stocks of different industry sector. One can note from Figure 2.19 that, excluding the crisis period, the

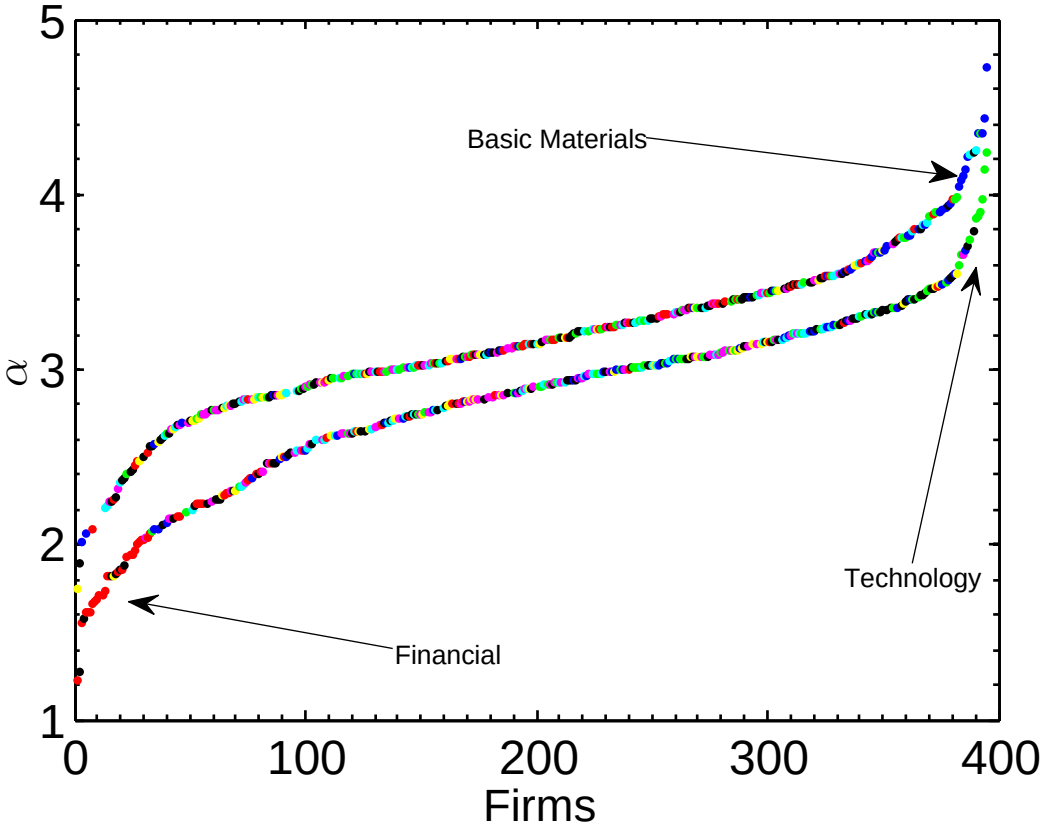

FIGURE 2.19: The tail exponents for all the companies analyzed including (lower curve) and excluding (upper curve) the time-period from December 2007 to April 2009, when the crisis occurred. We notice a clustering of the financial sector (red) at very low values of α, with many points lying in the region $\alpha < 2$. The other end of the curve, at high values of α, is instead mostly populated by the Technology (green), which has been the less affected by the crisis. This is in agreement with the renown fact that the financial sector was the one most profoundly affected by the crisis and whose fluctuations were the largest. Instead, before the crisis, the sector of Basic Material (blue) appears to be the most stable.

exponent increases across all firms and the occurrence of extreme events is much lower than the one observed when the crisis is included. In particular Figure 2.19

shows how the financial sector forms a cluster at the bottom left corner, when the crisis period is included, confirming that the most risky stocks over the crisis were those from the financial sector. It's also interesting to note that, throughout the crisis, some firms belonging to the Technology sector appear to be the most stable, whereas the more stable ones on the entire period are firms in the Basic Material sector.

The increase in the wGHE over the crisis reported in the previous section is therefore very likely to be influenced by the higher frequency of extreme returns in the same period. The Hurst exponent H is after all known to be related to the tail exponent for uniscaling processes. The relation which can be established is the following (see Appendix B)

$$H = \begin{cases} 1/\alpha & \text{if } \alpha < 2 \\ \\ 1/2 & \text{if } \alpha \geq 2 \, . \end{cases} \tag{2.23}$$

One can see that for fat tails, corresponding to $\alpha < 2$ the Hurst exponent H is expected to increase hyperbolically as α decreases. A very similar behaviour is predicted for multifractal time series, where one finds [15]

$$H(q) = \begin{cases} 1/\alpha & \text{if } q \leq \alpha \\ \\ 1/q & \text{if } q > \alpha \, . \end{cases} \tag{2.24}$$

Looking at $H(1)$ for fat tailed distribution means considering the case $q \leq \alpha$ and again the GHE is expected to increase as α decreases[13].

Tracking the dynamical evolution of the wGHE is thus a meaningful way of capturing the swings in complexity of financial time series. The wGHE has two advantages over the tail exponent:

1. It does not involve any arbitrariness in the choice of the range of values over which the fit is performed.

[13]The weighting procedure does not impact these theoretical relations.

2. It includes also information about the temporal structure of the time series, not only about the unconditional distribution.

The dynamical evolution of the wGHE therefore seems to be a useful tool to assess the riskiness of stocks over time. The significance of the observed fluctuations will be discussed in the next chapter.

2.5.4 Multifractality Changes and Risk

Let us further stress the relevance of the dynamical evolution of multifractality reported in this chapter. Multifractality changes in fact represent yet another tool to assess the riskiness and turmoil of the stage of the financial market. The intuition behind this claim is that periods of financial instability are typically characterised by an increase in the size of price swings - measured by the distribution's tails - which last for sizeable time lags - measured by volatility clustering. More generally, the complexity of the price increments time series increases over a financial crisis and this is reflected in the time variation of multifractality proxies. High degrees of multifractality are therefore synonymous of jitters and turbulent behaviour of the corresponding price and global increase in the average multifractality of a market should be tracked carefully as they correspond to an increase in the turbulence affecting the market. As we will see in Chapter 4 the 2008 financial crisis was characterised by an important increase in the average multifractality.

2.6 Summary

In this first chapter we have discussed the main empirical properties of stock returns, pointing out how these vary across different time scales. This has motivated us to focus the attention on the multifractal properties of financial time series and to discuss the major results known in the literature. In particular, we have stressed how multifractality can serve as an assessment tool for both the two major stock returns stylised facts. We have provided some further evidence of the

preponderance of fat-tails over long-range dependence in the contribution to multifractality, performing some original analysis aimed at identifying the different components contributing to the overall degree of multifractality. We have then introduced the wGHE and studied its dynamical evolution, which has revealed remarkable increasing trends for the firms bailed-out during the 2008 financial crisis. In agreement with previous works which have studied statically the heterogeneity of scaling exponents across the market, the observations reported in this chapter have led us to formulate the hypothesis that wGHE can be considered as a new tool to track the stability of a firm in time, where the degree of stability is inversely proportional to the absolute value of $H_t^w(q)$.

Chapter 3

Non-stationary Multifractality in Stock Returns

"Men occasionally stumble over the truth, but most of them pick themselves up and hurry off as if nothing ever happened."

- W. Churchill -

In this chapter we present a scaling analysis of equity returns aimed at unveiling non-stationarity in their multifractality. We compared empirical observations of the weighted generalised Hurst exponent (wGHE) with time series simulated via Multifractal Random Walk (MRW) in order to assess the validity of the constant multifractality hypothesis for stock returns. This chapter is based on the article *"Non-stationary Multifractality in Stock Returns"* [19].

To the best of our knowledge, tests for non-stationarity of multifractality in financial markets have never been investigated before in the Econophysics literature. The problem with the identification of significant dynamical fluctuations in multifractality proxies lies in their intrinsic noisy nature, which sometimes may let one think that a trend is present, whereas the dynamical variations are only due to the statistical uncertainty of the estimator used. In order to overcome this difficulty, one has to assume that the dynamics of asset prices is well-described by a particular model and then compare the entity of the observed fluctuations to those

envisaged by the selected model. In this chapter we follow this line of reasoning and among several multifractal models that have been proposed in the literature [28, 67–69], we choose to use as a benchmark the Multifractal Random Walk [20], because of its popularity in the Econophysics community as well as its structural features which make the comparison with the wGHE particularly straightforward. This chapter is organised as follows: in section 3.1 we introduce the general framework used to model asset returns in quantitative finance and we then discuss the MRW, the model which will be used to perform the hypothesis testing for non-stationary multifractality. Section 3.2 is devoted to the empirical analysis aimed at validating dynamical fluctuations in multifractality. A discussion is then proposed in Section 3.3.

3.1 Modelling Volatility: Multifractal Random Walk

The non-linear correlation exhibited by asset returns as well as the crucial role played by the powers of absolute values of returns in the multifractal analysis have suggested researchers to model returns as [26, 27]

$$r_{t,\tau} = \sigma_{t,\tau}\epsilon_t \tag{3.1}$$

where $\sigma_{t,\tau}$ is a strictly positive time correlated volatility process, which can carry dependence on the time scale τ, and ϵ_t are residuals, typically Gaussian, uncorrelated in time. In the following we will always drop the τ index unless necessary and refer to the volatility simply as σ_t and to returns as r_t. The decomposition in Equation (3.1) produces uncorrelated returns but ensures that for a proper choice of the volatility auto covariance function, non-linear correlations in returns are still possible. Note that one cannot actually observe the volatility process σ_t but only returns; for this reason, in order to give a quantitative measure of volatility, one needs to resort to *volatility proxies*, the most widely employed being the absolute returns $|r_t|$ and the square returns r_t^2. We plot in Figure 3.1 these two volatility

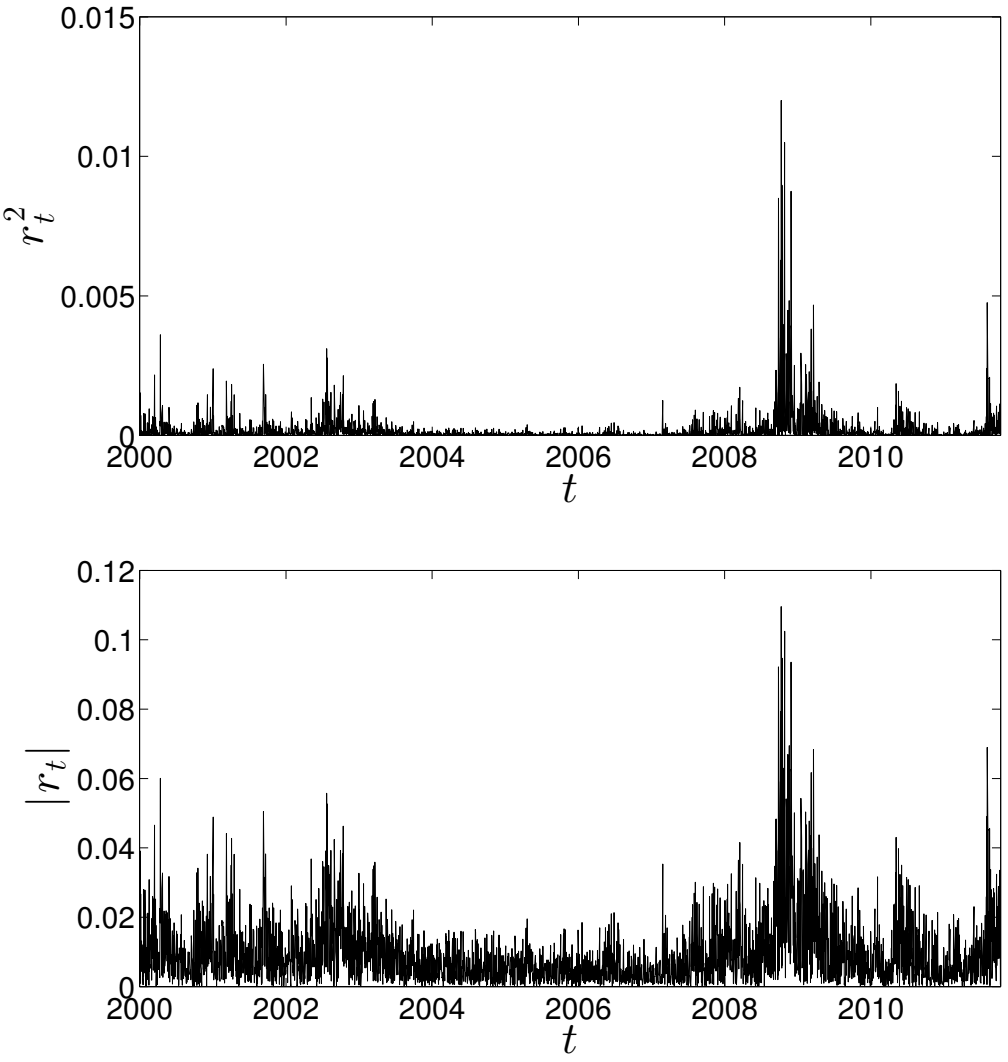

FIGURE 3.1: (Top) r_t^2 and $|r_t|$ (bottom) for the S&P 500 index in the period January 2000-January 2012.

proxies (square returns in the upper plot and absolute returns in the lower plot) for the S&P 500 index daily closing price in the period 2000-2012. One can see how both proxies reproduce very irregular profiles with unevenly spaced clusters of large fluctuations. These volatility bursts are responsible for the slow decay of the correlation functions $C_{|r_t|}(\tau)$ and $C_{r_t^2}(\tau)$ shown in Figure 2.7 in Chapter 2. Since the actual volatility is not observable, its definition is necessarily model dependent. In the literature there are mainly two branches of discrete time models tackling the problem of volatility modelling: GARCH models [86] and multifractal volatility models [87]. The first are conditional variance models, where the volatility process typically depends non-linearly on previous realisations of returns and

volatility itself through a feedback mechanism that produces non-trivial autocorrelation functions: details on the GARCH(p, q) models are reported in Appendix C. The second class instead models volatility as a multi-time scale process and long-range correlations arise from a cascade mechanism across time scales [88]. Aside from these two branches, many stochastic volatility models have been proposed in the literature both in discrete and continuous time [89–92].

Among multifractal models, the most popular one in the literature is probably the multifractal random walk (MRW) [20]. Differently from GARCH, its dynamics is not given conditionally but can be viewed as a stochastic volatility model constructed by taking the limit for $\Delta t \to 0$ of the process

$$X_{\Delta t}(t) = \sum_{k=1}^{t/\Delta t} \epsilon_{\Delta t}(k)\, e^{\omega_{\Delta t}(k)} \tag{3.2}$$

with $\epsilon_{\Delta t}(k)$ a Gaussian white noise with variance $\sigma^2 \Delta t$ and $e^{\omega_{\Delta t}(k)}$ a stochastic volatility uncorrelated with ϵ. The returns of Equation (3.1) for the MRW are therefore given by the process $\delta X_{\Delta t}(k) = \epsilon_{\Delta t}(k)\, e^{\omega_{\Delta t}(k)}$. By taking $\omega_{\Delta t}(k)$ as a stationary Gaussian process, we have log-normal volatility components. What distinguishes the limit $\Delta t \to 0$ of $X_{\Delta t}(t)$ from a Brownian motion is the choice of the auto covariance structure of the process $\omega_{\Delta t}(k)$, which is chosen, according to cascade-like processes [93], as

$$\mathrm{Cov}(\omega_{\Delta t}(k), \omega_{\Delta t}(k+h)) = \begin{cases} \lambda^2 \log \dfrac{T}{(1+h)\Delta t}, & h \leq T/\Delta t - 1 \\ 0, \quad \text{otherwise.} \end{cases} \tag{3.3}$$

The logarithmic decay with lag h of the auto covariance creates long memory in the process. This specification implicitly defines the integral scale T and the intermittency coefficient λ. In order for the process $X_{\Delta t}$ to have finite variance in the limit $\Delta t \to 0$, one should impose $E(\omega_{\Delta t}(k)) = -\mathrm{Var}(\omega_{\Delta t}(k)) = -\lambda^2 \ln(T/\Delta t)$ [20]. $X(t)$ can be shown to obey self similarity exactly, i.e. for a time scale contraction $\tau' = s\tau$ $(s < 1)$

$$X(t + \tau') - X(t) = e^{\Gamma_s}[X(t+\tau) - X(t)], \tag{3.4}$$

with Γ_s a Gaussian random variable of variance $\mathrm{Var}(\Gamma_s) = -\lambda^2 \ln(s)$ [94]. From equation (3.4) one can show that [95]

$$\mathbb{E}(|X(t+\tau) - X(t)|^q) = K_q \left(\frac{\tau}{T}\right)^{\zeta_q}, \tag{3.5}$$

with K_q a q-dependent factor and ζ_q a non linear function of q, i.e. the process $X(t)$ is multi-scaling. It should be understood that the scaling (3.5) is exact only in the continuous time limit $\Delta t \to 0$. Nonetheless for $\Delta t < \tau$ one can recover good approximations of the scaling even when considering the discretized version $X_{\Delta t}(t)$.

The main appeal of this model for describing stock returns evolution lies in its ability to reproduce faithfully the most common stylised facts of financial markets: the hyperbolic decay of the volatility auto covariance function (3.3) as well as the heavy tails of the process increments. Indeed, as shown in [94], the probability of observing increments larger than a certain value x decays as a power law for large x:

$$\mathbb{P}(|X(t+\tau) - X(t)| > x) \sim x^{-\frac{1}{2\lambda^2}}. \tag{3.6}$$

The parameter λ controls therefore also the thickness of the tails of the returns distribution. It must however be noted that for the usual values found for λ on empirical data ($\lambda^2 \sim 0.03$), the model produces power-law tails whose exponent $\alpha = 1/2\lambda^2$ is almost ten times larger than that observed for financial returns [94] (whose α is notoriously in the range [2,5] [27]) and therefore is out of line with empirical observations. The stationarity and the causal structure of this model though make it also preferable to other multifractal models [96]. One of the most important features is that the multi-fractal spectrum of this model can be computed exactly to be [20]

$$\zeta_q = (q - q(q-2)\lambda^2)/2. \tag{3.7}$$

The scaling function ζ_q is therefore a parabola whose constant concavity only depends on λ^2. We show in Figure 3.2 a realisation of the MRW (left) along with its increments (right), simulated with parameters $\lambda = 0.2, T = 500$ and $\sigma = 0.9$:

as it will be shown in the following section, these values are very similar to those obtained on empirical data.

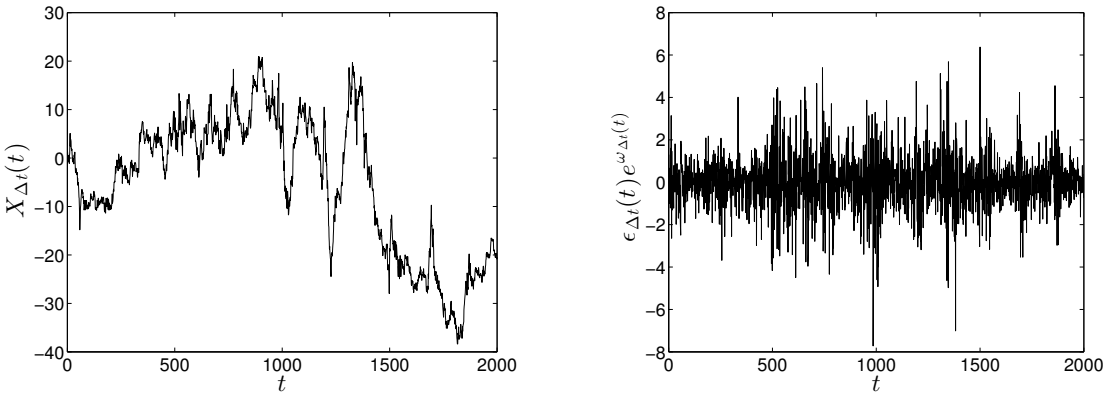

FIGURE 3.2: We show a realisation of MRW process (left) and its corresponding increments (right) for a simulation with parameters $\lambda = 0.2$, $T = 500$ and $\sigma = 0.90$.

It has been shown that modifications of the log-normal model, where residuals $\epsilon_{\Delta t}(t)$ are Gaussian, can account better for the fat tails observed in empirical data [72]. One may first consider residuals $\epsilon_{\Delta t}(t)$ to be Student-t distributed, that is (we drop the subscript Δt for readability)

$$P(\epsilon) = \frac{1}{\sqrt{\pi}} \frac{\Gamma\left(\frac{1+\nu}{2}\right)}{\Gamma(\frac{\nu}{2})} \frac{a^{\nu}}{(\epsilon^2 + a^2)^{\frac{1+\nu}{2}}}, \tag{3.8}$$

where ν is the number of degrees of freedom, Γ is the Euler gamma function and the parameter a is related to the variance σ^2 via $\sigma^2 = a^2/(\nu - 2)$ [27]. The Student-t distribution is known to reproduce well stock returns tails, if one considers $\nu = 4$ or 5 [27]. In order to let the model have fatter tails one can also act on the unconditional distribution of the $\omega_{\Delta t}$; in [72] indeed, it was pointed out that a cascade model where $\omega_{\Delta t}$ follows a gamma law can account for empirical observations better than a normal law. The pdf of $\omega_{\Delta t}$ is given in this case by (again the subscript Δt is dropped)

$$P(\omega) = \frac{1}{\Gamma(k)\theta^k} \omega^{k-1} e^{-\omega/\theta}, \tag{3.9}$$

where k and θ are respectively shape and scale parameters.

3.2 Time-varying Multifractality

3.2.1 Intermittency and GHE

Let us here start to look at the theoretical relation between the intermittency coefficient λ and the GHE $H(q)$ that can be established from the identity $\zeta_q = qH(q)$, introduced in Section 2.3.1 in Chapter 2. Since the scaling exponents $H(q)$ are computed from the scaling of the empirical moments, one has for the log-normal MRW that q is bounded from above by [72]

$$q \leq \frac{\sqrt{2}}{\lambda}. \tag{3.10}$$

Hence λ and q must be chosen jointly in such a way that the last inequality holds. As shown in Section 3.2.3, typical values of λ obtained for financial data analysed in this study very rarely exceed the value of 0.3, which is well within the bound of equation (3.10). By considering $q = 1, 2, 3$ the scaling relation $\zeta_q = qH(q)$ for the MRW (using equation (3.7)) reads

$$\begin{aligned} H(1) =& \frac{1}{2}(1 + \lambda^2), \\ H(2) =& \frac{1}{2}, \\ H(3) =& \frac{1}{2}(1 - \lambda^2). \end{aligned} \tag{3.11}$$

Values of the expressions (3.11) for different λ's are reported in Table 3.1 for $\lambda = 0.2, 0.25, 0.3$. In order to compare theoretical predictions with GHE's computed on synthetic time series, we have performed the following statistical study : for a given value of λ, we have simulated 1000 MRW series and computed the corresponding GHE's for every synthetic series; then, assuming these GHE's are i.i.d., which is quite reasonable since every MRW simulation is independent from the others, we have obtained a sample distribution for the GHE's. From this distribution we have computed the $\{2.5\%, 50\%, 97.5\%\}$-quantiles which give the range of fluctuation

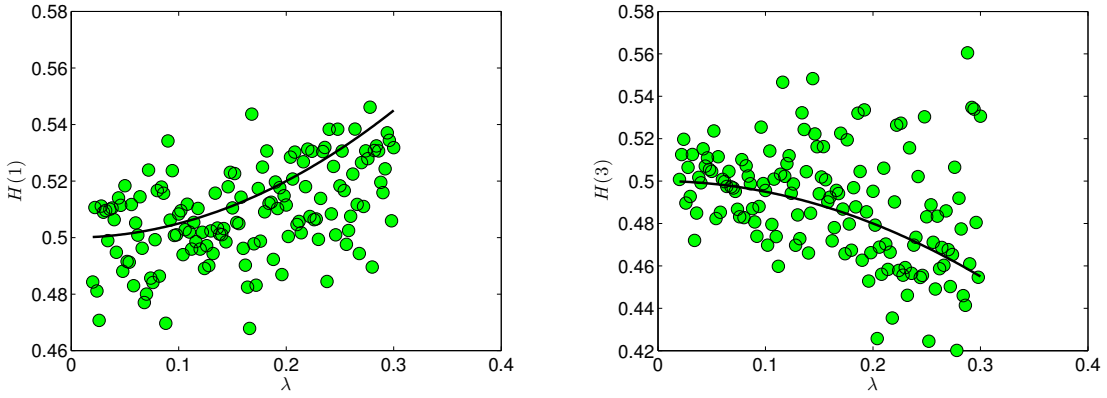

FIGURE 3.3: We plot the $H(q)$ vs λ for $q = 1$ (left) and $q = 3$ (right). The dots refer to the H(q)'s values obtained from the MRW-simulated time series with λ ranging in [0.01,0.3], $T = 500$ and $\sigma = 1$. The black solid lines are the theoretical relations (3.11) for $q = 1, 3$. The GHE have been obtained with τ_{max} varying between 10 and 30, whereas Δt has been set to $\Delta t = 1$.

of the GHE beyond which any observation can be deemed anomalous. We have performed this study for $q = \{1, 2, 3\}$, after having checked that the scaling doesn't hold for larger q's. Results of this analysis for $\lambda = 0.2, 0.25, 0.3$ and $T = 1250, 2500$ are reported in Table 3.2. By comparing these values with those expected from the MRW model in Table 3.1 we can conclude that all values computed via equations (3.11) are compatible with the statistical confidence intervals computed on the synthetic time series. In Figure 3.3 we plot the scaling exponents $H(1)$ and $H(3)$ computed on synthetic MRW time series with intermittency λ ranging in $[0.01, 0.3]$ together with the theoretical relations (3.11) (black solid lines). The measured scaling exponents are found to be fluctuating around the theoretical relations. Although in this figure we show only the cases $q = 1, 3$, we have also verified that the case $q = 2$ is consistent with the expected behaviour $H(2) = 1/2$.

TABLE 3.1: Values of $H(q)$ for $q = \{1, 2, 3\}$ computed through the expressions in equation (3.11) for $\lambda = 0.2, 0.25, 0.3$.

	$\lambda = 0.2$	$\lambda = 0.25$	$\lambda = 0.3$
H(1)	0.52	0.53	0.54
H(2)	0.50	0.50	0.50
H(3)	0.48	0.47	0.45

TABLE 3.2: We report numerical results of the $\{2.5\%, 50\%, 97.5\%\}$-quantiles obtained from the distributions of the GHE's computed on several synthetic MRW time series for $q = 1, 2, 3$, for $\lambda = 0.2, 0.25, 0.3$ and $T = 1250, 2500$. We can appreciate dependence of the quantiles both on λ and q.

		$\{2.5\%, 50\%, 97.5\%\} - \text{quantiles}$		
		H(1)	H(2)	H(3)
$\lambda = 0.2$	T=1250	$\{0.4782, 0.5110, 0.5427\}$	$\{0.4592, 0.4986, 0.5381\}$	$\{0.4326, 0.4838, 0.5421\}$
	T=2500	$\{0.4768, 0.5101, 0.5450\}$	$\{0.4562, 0.4965, 0.5407\}$	$\{0.4284, 0.4807, 0.5454\}$
$\lambda = 0.25$	T=1250	$\{0.4844, 0.5167, 0.5502\}$	$\{0.4527, 0.4976, 0.5415\}$	$\{0.4029, 0.4739, 0.5488\}$
	T=2500	$\{0.4843, 0.5180, 0.5517\}$	$\{0.4530, 0.4988, 0.5428\}$	$\{0.4038, 0.4774, 0.5479\}$
$\lambda = 0.3$	T=1250	$\{0.4900, 0.5249, 0.5616\}$	$\{0.4374, 0.4962, 0.5571\}$	$\{0.3642, 0.4580, 0.5597\}$
	T=2500	$\{0.4878, 0.5253, 0.5626\}$	$\{0.4398, 0.4968, 0.5558\}$	$\{0.3743, 0.4653, 0.5609\}$

3.2.2 Statistical Testing Procedure

The hypothesis we want to test in this section is whether stock returns exhibit multifractal properties consistent with a multifractal model whose intermittency coefficient λ is constant. To this end, we have compared dynamical fluctuations of the multifractality proxy $\Delta H^w(1,2)$ for empirical time series with fluctuations observed on synthetic MRW time series simulated with parameters calibrated on the corresponding empirical time series. More specifically, we have performed the following statistical testing procedure: after having estimated the MRW parameters from the empirical data we have simulated 1000 MRW synthetic series of length $n = 1250$ with the parameters obtained from the empirical time series. On each of these simulated time series the scaling exponents $H^w(1)$ and $H^w(2)$ have been computed as average of several fits of the relation $\log M_q(t,\tau) \sim qH^w(q)\log\tau$ with the scale $\tau \in [1, \tau_{max}]$, with τ_{max} varied between 10 and 30. The moments $M_q(t,\tau)$ are estimated as $\langle |r_{t,\tau}|^q \rangle_w \sim \sum_{s=0}^{\Delta t-1} w_s |r_{t,\tau}|^q$, where $w_s = w_0 \exp\left(-\frac{s}{\theta}\right)$, $\forall s = 1, \ldots, \Delta t - 1$, with θ being the damping coefficient, set equal to 415 days, in agreement with previous analysis carried out in Chapter 2. The multifractality proxy has then been computed as $\Delta H^w(1,2) = H^w(1) - H^w(2)$. Repeating this procedure for each of the simulated series we have obtained a distribution of the measured $\Delta H^w(1,2)$, whose width is a measure of the range of fluctuation of $\Delta H^w(1,2)$. We have then computed dynamically on 1250 days rolling windows, $\Delta H_t^w(1,2)$ on the empirical time series and compared its fluctuations in time with the extreme quantiles of the distributions obtained from synthetic series. We have then looked at the percentage of fluctuations falling outside of the confidence intervals provided by the extreme quantiles. A high quantile-exceedance rate of the empirical $\Delta H_t^w(1,2)$ would suggest that the fluctuation in multifractality envisaged by the model are smaller than those observed on empirical data. If this is the case, these findings would suggest that a time-varying intermittency coefficient may be responsible for this feature observed in financial time series.

This statistical test has been performed on a set of 342 stocks daily closing prices quoted in the NYSE in the period ranging from 01-01-1995 to 22-10-2012.

3.2.3 Results

As discussed in the previous section, the first step of the testing procedure is obtaining the MRW parameters from each empirical time series. The parameters of the MRW have been extracted from the empirical data through the behaviour of the log-volatility auto covariance function. By identifying empirical daily returns with the increments of the MRW, using Equation (3.2) one can write

$$\log(|r_{t,1}|) = \log(|\delta X_{\Delta t}(k)|) = \omega_{\Delta t}(k) + \log|\epsilon_{\Delta t}(k)|. \tag{3.12}$$

It has been shown [88] that for financial time series, the function

$$C(h) = \mathrm{Cov}(\log|r_{t+h,\tau}|, \log|r_{t,\tau}|) \tag{3.13}$$

exhibits a slow decay with the lag h. Note that for daily returns the logarithm of the absolute values is always well defined, as the price increments are always different from zero. By putting together the last two equations and using Equation (3.3) we see that $C(h)$ is proportional to the log-volatility auto covariance function plus a term uncorrelated in the lag h, i.e.

$$C(h) \sim \lambda^2 \log \frac{T}{(1+h)\Delta t}. \tag{3.14}$$

The intermittency coefficient λ and the integral scale T can therefore be estimated from linear fits of equation (3.14) in log-linear scale. The parameter σ is instead estimated through the relationship $\mathrm{Var}(\delta X_{\Delta t}(k)) = \mathrm{Var}(r_{t,1}) = \sigma^2$. As already established in previous studies [95], the estimation of λ through this method yields much more reliable results than an estimation based on the variogram of the $\omega_{\Delta t}(k)$ as a function of the scale Δt (see [95] for details and comparison of the two methods). In the whole procedure we fix the return scale at $\Delta t = 1$ day. These simulations of MRW synthetic series have been performed using a Fast Fourier Transform of the auto covariance function in equation (3.3).

TABLE 3.3: In this table we report, for 13 different companies, the parameters λ, T, σ of the MRW obtained from empirical data. The fourth column shows the $\{2.5\%, 50\%, 97.5\%\}$-quantiles of the distributions of $\Delta H^w(1,2)$ computed from 1000 realisations of MRW synthetic series of length 1250, simulated with the corresponding set of parameters. The fifth column shows the percentage of dynamical $\Delta H^w(1,2)$'s, computed on 1250 time steps rolling windows on the empirical time series, falling outside the confidence interval given in the fourth column.

	λ	T (days)	σ	Quantiles	Exceedances percentage
Boeing Corp	0.1228	1260	1.0651	$\{-0.0205, 0.0039, 0.0265\}$	8 %
Microsoft Corp	0.1342	739	0.021	$\{-0.03341, 0.01067, 0.04675\}$	33 %
PNC Financials	0.1418	28201	0.025	$\{-0.03191, 0.0098, 0.0349\}$	7 %
Sara Lee Corp	0.1335	3746	0.019	$\{-0.03381, 0.00948, 0.04524\}$	31.5%
Rowan Cos INC	0.2046	1080	0.7992	$\{-0.0215, 0.0066, 0.0296\}$	11.4%
IBM Corp	0.2363	816	1.7811	$\{-0.0214, 0.0068, 0.0310\}$	14%
Wells Fargo	0.3010	631	0.5923	$\{-0.0237, 0.0134, 0.0433\}$	15%
American Express	0.1340	1520	0.8356	$\{-0.0212, 0.0125, 0.0429\}$	29%
General Motors Corp	0.1727	440	1.0291	$\{-0.0211, 0.0063, 0.0278\}$	41%
Citigroup	0.2196	568	6.4585	$\{-0.0202, 0.0146, 0.0458\}$	39%
JPMorgan Chase	0.2950	534	0.9792	$\{-0.0219, 0.0097, 0.0356\}$	27 %
Dominion Resources INC	0.1738	754	0.013	$\{-0.0317, 0.01372, 0.05411\}$	10 %
Morgan Stanley	0.2528	538	1.1185	$\{-0.0221, 0.0079, 0.0333\}$	14 %

We have estimated the MRW parameters for each empirical time series. Then, as outlined in Section 3.2.2, for each set of parameters we have simulated 1000 MRW synthetic series of 1250 time steps and on each series computed $\Delta H^w(1,2)$. The distribution of these $\Delta H^w(1,2)$ for each set of parameters is used to estimate the $\{2.5\%, 50\%, 97.5\%\}$-quantiles. Values of the parameters obtained for different stocks together with the corresponding quantiles obtained from the distributions of the $\Delta H^w(1,2)$ obtained from the simulated MRW series are reported in Table 3.3. In Table 3.3 we also report the percentage of empirical $\Delta H_t^w(1,2)$ computed on rolling windows of length 1250 exceeding the quantiles, defined as

$$\text{Exceedance} = \frac{1}{N}\left(\sum_{t=1}^{N} \mathbb{1}_{\{\Delta H_t^w(1,2) > Q^w_{97.5\%}\}} + \sum_{t=1}^{N} \mathbb{1}_{\{\Delta H_t^w(1,2) < Q^w_{2.5\%}\}}\right), \tag{3.15}$$

where N is the number of time windows, $Q^w_{97.5\%}$ and $Q^w_{2.5\%}$ are respectively the 97.5% and 2.5% quantiles of the $\Delta H^w(1,2)$ distribution and $\mathbb{1}_{\{x\}}$ is the indicator function which is 1 if the condition x is enforced and 0 otherwise. The rather high number of points falling outside the confidence intervals confirms systematically that empirical data do not agree with the hypothesis of constant multifractality. As one can appreciate from the quantiles reported in Table 3.3, there is no correlation between the number of quantile-exceeding $\Delta H_t^w(1,2)$ and λ. This fact tells us that a simple underestimation of λ is unlikely to be the cause of the high quantile exceedance rate. In Figure 3.4 we show results of this study for several daily stock prices, where $\Delta H_t^w(1,2)$ is shown to exceed the quantiles many times.

We must remark that the values obtained for the correlation lengths T are very large compared to the time series lengths, a feature which is nonetheless commonly observed in multi-timescale volatility models [97]. However, multifractality is expected to be proportional only to λ^2 and thus the values of $\Delta H^w(1,2)$ that we have measured should not be affected by the true T of the series. This has been checked by computing the quantiles of the distributions of $\Delta H^w(1,2)$ simulated with a fixed λ and varying the integral scale T. As one can appreciate in Figure 3.5 the quantiles do not vary when the integral scale of the simulated series is changed but λ is kept constant.

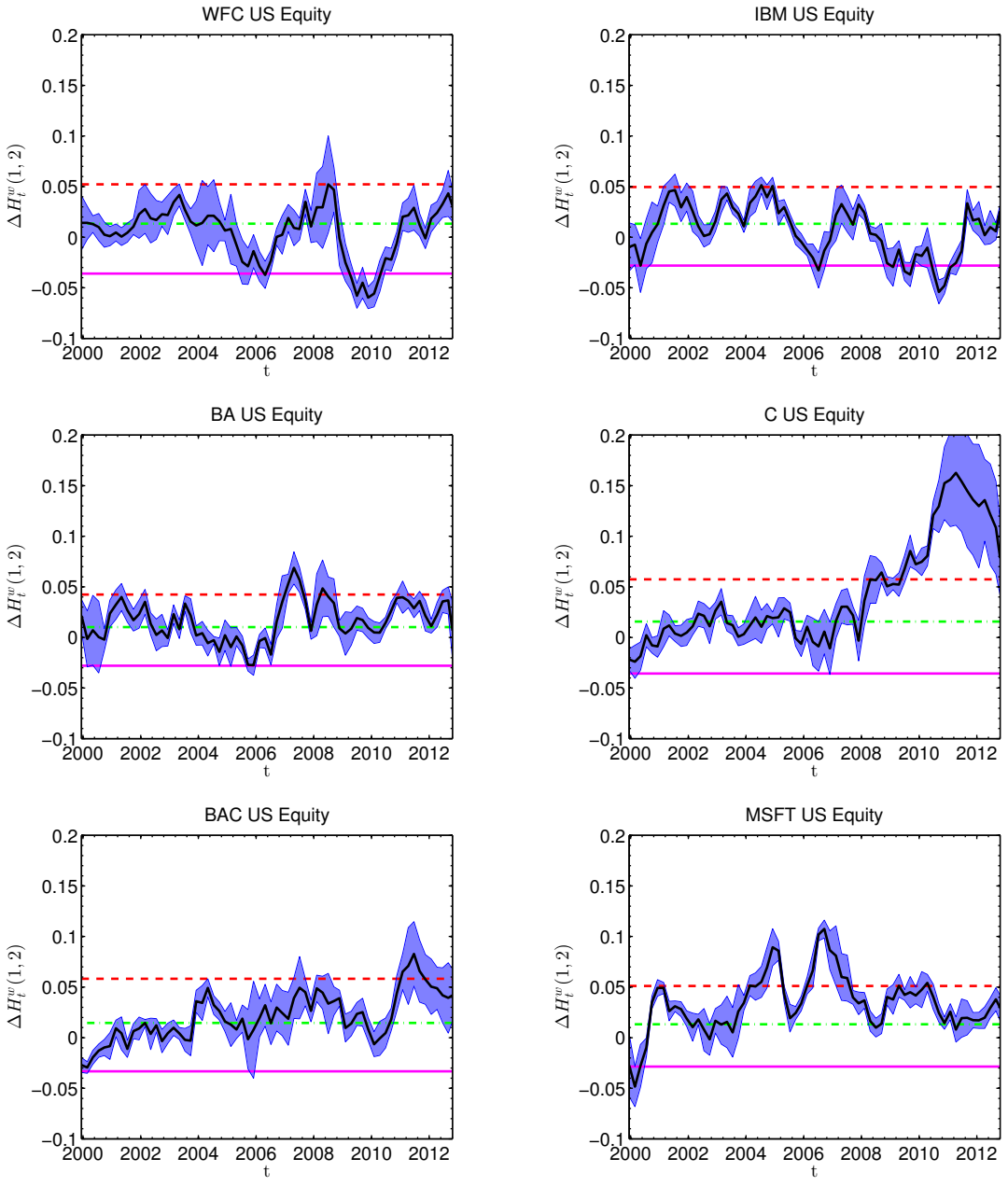

FIGURE 3.4: *We plot the dynamical evolution $\Delta H_t^w(1,2)$ (black thick line) for six stocks in the time period 01-01-1995 to 22-10-2012. $\Delta H_t^w(1,2)$ is evaluated on moving overlapping time windows of 1250 days, with a shift of 100 days. The first time window includes data from 1995 to 2000 and the corresponding $\Delta H_t^w(1,2)$ is plotted at the end of the period. The horizontal lines represent the 2.5% (red dashed line), 50% (green dot-dashed line) and 97.5% (magenta continuous line) quantiles extracted from the distribution of $\Delta H^w(1,2)$ estimated from 1000 realisations of MRW with parameters extracted from the data. The blue shaded area represents the error on $\Delta H_t^w(1,2)$, computed, for each t, as the standard deviation on $\Delta H^w(1,2)$.*

We have performed this statistical test for the set of 342 stocks listed in the NYSE. In particular, for each empirical time series, we have compared the fluctuations

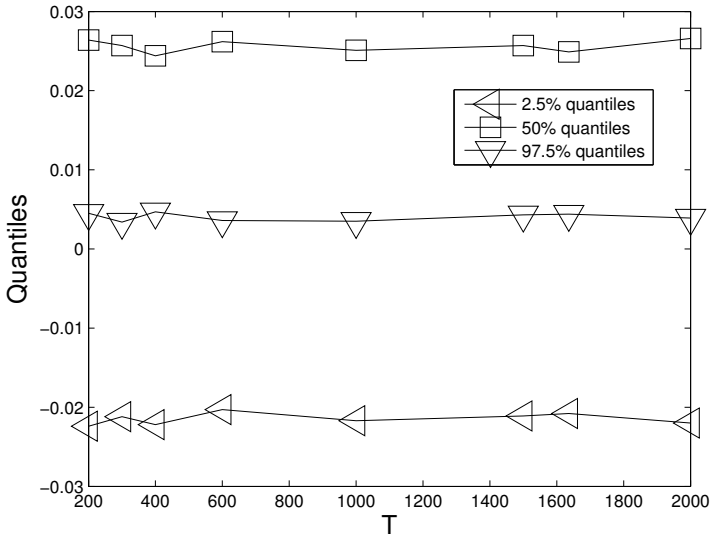

FIGURE 3.5: *We plot the quantiles of the* $\Delta H^w(1,2)$ *distribution as function of the integral scale* T*. We have fixed* $\lambda = 0.2$ *and computed the* $\Delta H^w(1,2)$ *quantiles on synthetic series simulated with integral scales* $T = 200, 300, 400, 600, 1000, 1500, 1650, 2000$*.*

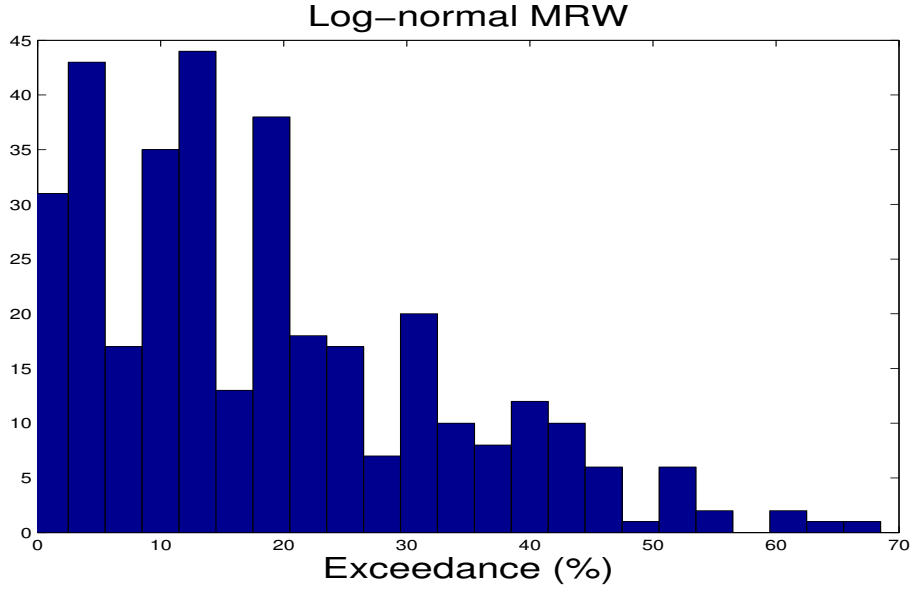

FIGURE 3.6: *Histogram of the observed exceedance percentages observed on the whole set of stocks when the benchmark model is the log-normal MRW. On the y-axes is the number of stocks exhibiting the percentage of quantile-crossings (given in the x-axis).*

of empirical $\Delta H_t^w(1,2)$ with the confidence intervals obtained from the MRW series simulated with the parameters estimated from the corresponding stock. The histogram in Figure 3.6 shows the distribution of the observed percentage of exceedances on this data set. It is remarkable that a conspicuous number of stocks shows a relevant number of $\Delta H_t^w(1,2)$'s falling outside the 95% confidence interval: the majority of stocks exhibits quantile crossings above 10% and we find even some percentages above 60%. This histogram shows that globally the hypothesis of a log-normal MRW model with constant intermittency appears to be a poor approximation of empirical time series behaviour.

The conclusion one draws from these observations is that our proxy of multifractality seems to defy the hypothesis of a constant multifractal behaviour, as in the setup of the MRW model, as the majority of stocks show time variations of $\Delta H^w(1,2)$ which are not in agreement with the null hypothesis of constant multifractal behaviour. As discussed above, a time varying multifractality requires a time varying λ. The observations reported here suggest that the scaling properties of financial time series may vary over time because of the reflection of complex and varying economic constraints. However, one needs to pay attention to the fact that the log-normal MRW with Gaussian residual may be a poor approximation for the empirical observations of the return distribution tails. In the next two sections we perform the same analysis considering two variations of the model with thicker tails.

3.2.4 The effect of Student-t residuals

We have carried out the same analysis performed in the previous section for the complete set of 342 stocks, this time taking the residuals ϵ to be Student-t distributed with $\nu = 4$ degrees of freedom, a value which is known to reproduce well financial returns [27]. MRW time series with Student-t residuals (tMRW) have tails fatter than those observed in the log-normal case. We have simulated 1000 realisations of tMRW and computed the proxy of multifractality $\Delta H^w(1,2)$ along with the corresponding $\{2.5\%, 50\%, 97.5\%\}$-quantiles. We plot in Figure 3.7 the

observed $\Delta H_t^w(1,2)$ for a selection of stocks together with the quantiles of the distribution of $\Delta H^w(1,2)$ computed from the 1000 tMRW simulations. We also show in the top plot of Figure 3.10 the observed percentage of quantile-overpassing for all stocks analysed: as this plot shows, we still retrieve a significant fraction of the observations overpassing the extreme quantiles, with many $\Delta H^w(1,2)$ still overpassing the quantiles more than 10% of the time. The rate of quantile-crossing is nonetheless reduced with respect to the case in which residuals are Gaussian. The overall multifractal properties of the simulated time series are very much influenced by the Student-t residuals: we observe indeed a systematic shift upwards of the sample distributions of the $\Delta H^w(1,2)$. As shown in Figure 3.8, the median, corresponding to the 50% quantile, is now peaked around 0.04, whereas the median retrieved from the log-normal case is peaked around 0.015. This tells us that, for a fixed intermittency, the thickness of the tails dramatically rebounds on the multifractal properties of the simulated series, which have now a higher degree of multifractality. This effect, which has also been reported in [17], clearly distinguishes different components of the measured multifractal property of the process: on the one hand the non-linear temporal dependence of the series, on the other hand the thickness of the tails of the unconditional distribution. Both aspects contribute to the measured multifractality and, by thickening the tails of the returns distribution, one can account for most of the anomalous fluctuations observed in the scaling exponents. Nonetheless the remaining multifractality still appears to defy the hypothesis of constant volatility covariance. In other words, the presence of more extreme fluctuations in the return process cannot fully account for the anomalous fluctuations observed in the empirical $\Delta H_t^w(1,2)$.

3.2.5 The effect of log-gamma volatility

Another possible modification of the standard log-normal MRW in order to account for fatter tails is to consider the volatility to be log-gamma distributed. It has been shown in [72] that such a specification can reproduce the fat tails observed in the empirical distribution of stock returns better than the log-normal MRW.

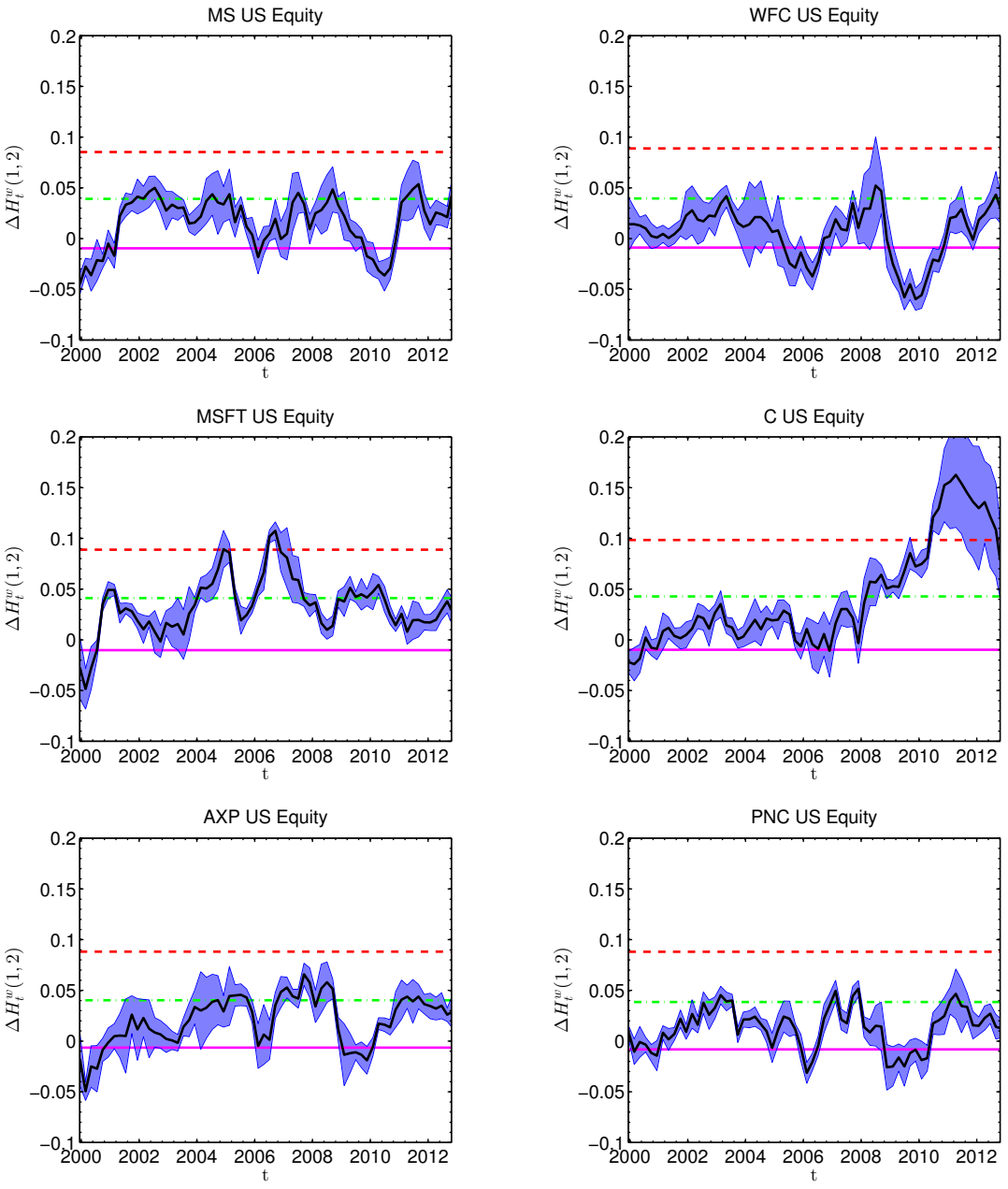

FIGURE 3.7: *We plot the dynamical evolution* $\Delta H_t^w(1,2)$ *(black thick line) for the six stocks in the time period 01-01-1995 to 22-10-2012. The horizontal lines represent the 2.5% (red dashed line), 50% (green dot-dashed line) and 97.5% (magenta continuous line) quantiles extracted from the distribution of* $\Delta H^w(1,2)$*, obtained from many simulations of MRW with Student-t distributed residuals. The blue shaded area represents the error on* $\Delta H_t^w(1,2)$*, computed, for each t, as the standard deviation on* $\Delta H^w(1,2)$*. The distributions of the simulated returns have thicker tails than in the original log-normal case, and this results in the quantiles of the sample distributions of* $\Delta H^w(1,2)$ *being larger.*

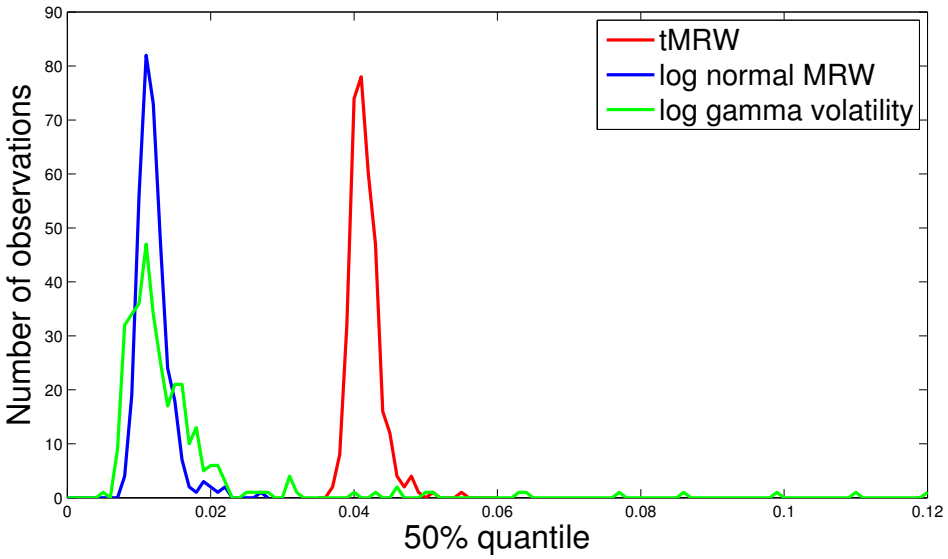

FIGURE 3.8: *(Color Online) We plot the sample distributions of the* 50% *quantiles observed on 1000 simulations of the three different specifications of the MRW with the parameters extracted from empirical data: log-normal MRW (blue line), MRW with Student-t residuals (red line) and MRW with log-gamma volatility (green line). On the x-axis is the value of the observed quantile, while on the y-axis we report the number of observations.*

Synthetic time series with log-gamma volatility components are obtained replacing the normal distribution with the law given in (3.9) for ω. We have repeated the same testing procedure: after having estimated the model parameters we have computed the $\Delta H^w(1,2)$ confidence intervals and, for each stock, we have compared the fluctuations of $\Delta H_t^w(1,2)$ with the corresponding confidence interval. Examples of $\Delta H_t^w(1,2)$ for a set of stocks compared with quantiles obtained with this specification of the volatility are shown in Figure 3.9. We also show in the histogram at the bottom of Figure 3.10 the percentage of quantile-crossing observed on the complete data set. The histogram shows a dramatic reduction of quantile exceedances, with most of the anomalous fluctuations now falling below 2% and therefore in agreement with what envisaged by the log-gamma MRW. As shown in Figure 3.8, the median of the $\Delta H^w(1,2)$ distribution, although showing some large values, is distributed around the same values observed for the log-normal MRW. What is remarkably more broadly distributed than in the two previous cases is the 97.5% quantile, whose observed distribution on the whole data set of 342 stocks is shown in Figure 3.11, compared with those retrieved for log-normal MRW and

tMRW. This confirms that the log-gamma volatility provides a framework in which very large fluctuations of the measured multifractality are much more likely than in the other two cases inspected but, as Figure 3.10 shows, some stocks still show fluctuations of $\Delta H^w(1,2)$ which not even the log-gamma MRW can explain.

Overall we can say that both modifications of the log-normal MRW model confirm that beefing up the tails of the distribution of synthetic time series has a sizeable impact on multifractal properties, with the percentage of anomalous fluctuations being drastically reduced. Nonetheless, we still observe some stocks systematically overpassing the confidence intervals that do not allow to accept globally the hypothesis of constant multifractality.
As a further test of non-stationarity we have performed Kwiatkowski-Philipps-Schmidt-Shin (KPSS) test and Leybourne-McCabe [98] test on the time series of $\Delta H_t^w(1,2)$ [1]. Both tests assess the hypothesis that the univariate time series of $\Delta H_t^w(1,2)$ are trend stationary against the alternative of non-stationary process. For each stock we have performed the two tests on the $\Delta H_t^w(1,2)$ time series. Since $\Delta H_t^w(1,2)$ is computed on the moving window, the time series have length 57. Both tests clearly confirm that the majority of stocks do not show stationary multifractal properties: the hypothesis of stationarity is rejected for 312 stocks (out of the total of 342) in the KPSS test and for 230 for the Leybourne-McCabe test, in both cases with a p-Value $p = 0.01$. This further confirms the non-stationarity of the multifractal indicator.

3.3 Discussion

The main scope of the research carried out in this chapter has been to measure and validate variations of multifractality on a set of stock returns. As a benchmark we have considered the MRW model, a parsimonious multifractal model that is able to reproduce faithfully many features commonly observed in stock returns. Since the dynamical estimations of multifractality are made on relatively small samples,

[1] Both tests have been performed by means of the MATLAB routine functions "kpsstest" and "lmctest".

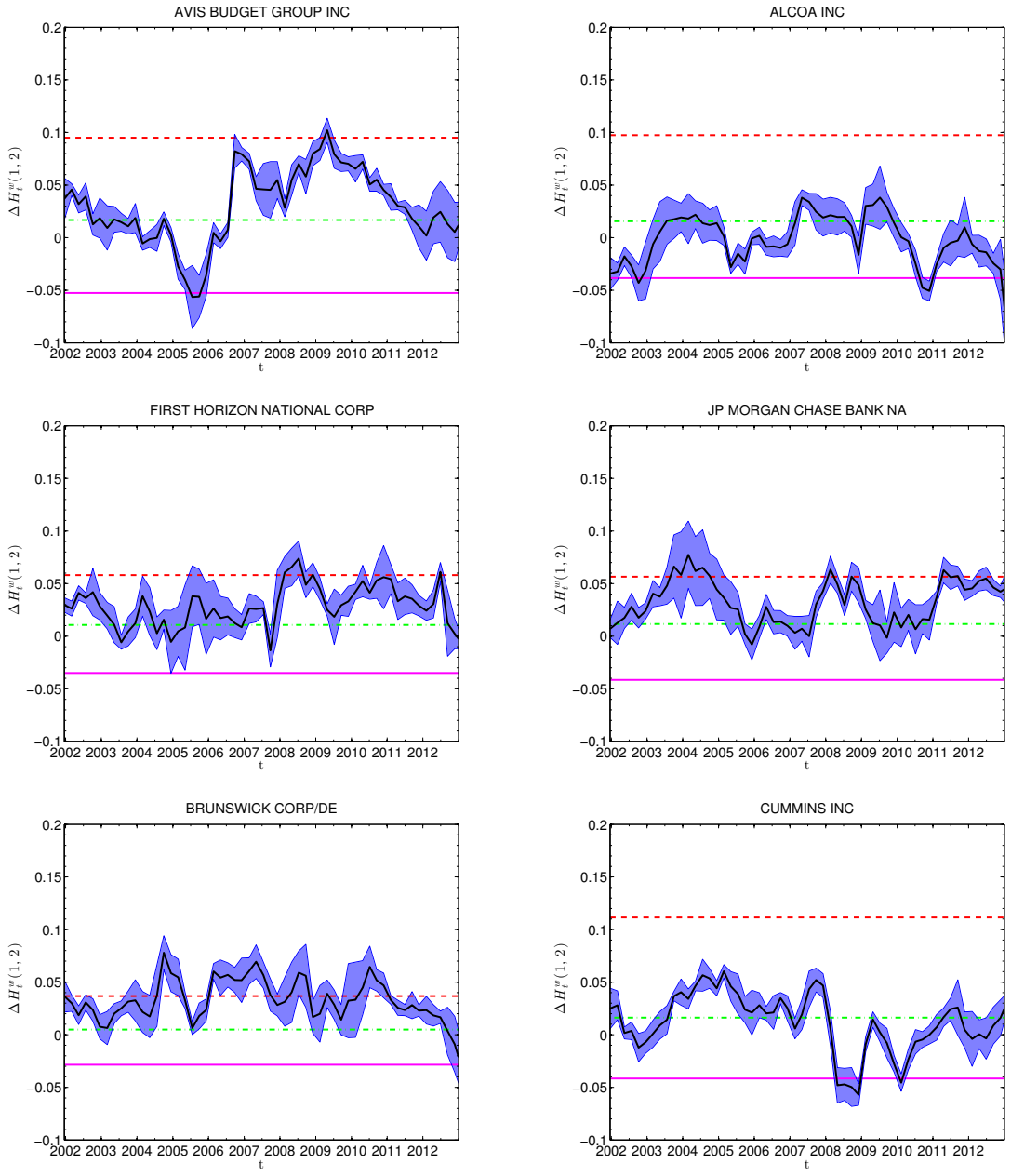

FIGURE 3.9: *We plot the dynamical evolution* $\Delta H_t^w(1,2)$ *(black thick line) for six stocks in the time period 01-01-1995 to 22-10-2012. The horizontal lines represent the 2.5% (red dashed line), 50% (green dot-dashed line) and 97.5% (magenta continuous line) quantiles extracted from the distribution of* $\Delta H^w(1,2)$*, obtained from many simulations of MRW with the volatility process* ω *simulated as a gamma random variable with* $k = \theta = 1$*. The blue shaded area represents the error on* $\Delta H_t^w(1,2)$*, computed, for each* t*, as the standard deviation on* $\Delta H^w(1,2)$*.*

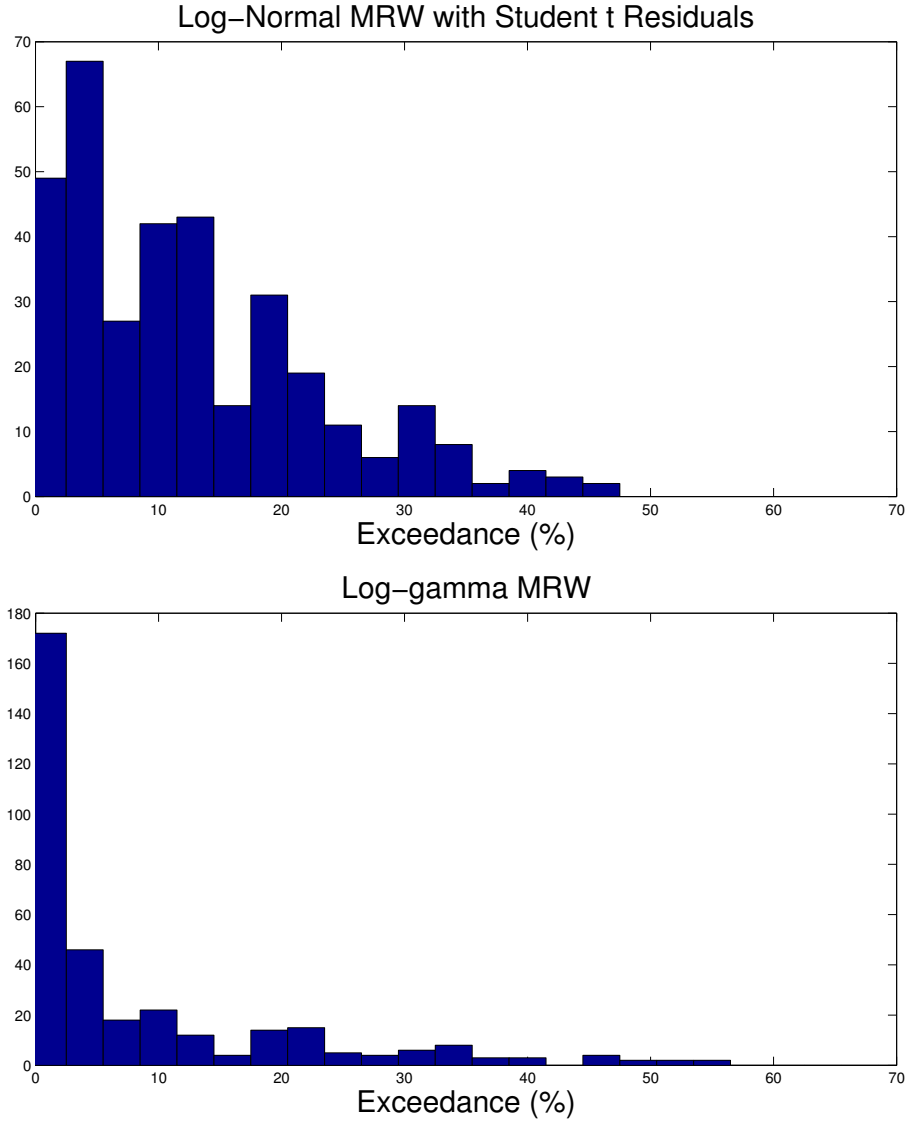

FIGURE 3.10: *Histogram of the observed exceedance percentages of the whole set of stocks when the benchmark model is taken to be the log-normal MRW with Student-t residuals (top) and the log-gamma MRW (bottom). On the y-axes is the number of stocks exhibiting the percentage of quantile-crossings given in the x-axis. The reduction of the overall number of exceedances is remarkable if compared with that obtained for log-normal MRW, whose histogram is shown in Figure 3.6.*

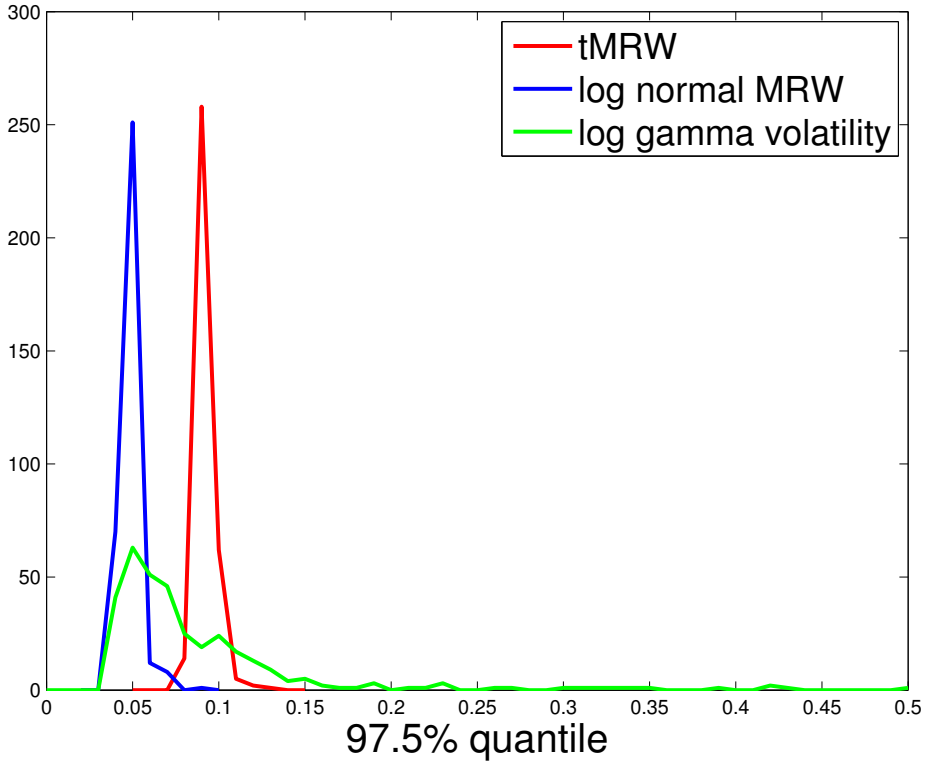

FIGURE 3.11: *We plot the sample distributions of the* 97.5% *quantiles extracted from 1000 simulations of the three different specifications of the MRW with the parameters extracted from each empirical time series: log-normal MRW (blue line), MRW with Student-t residuals (red line) and MRW with log-gamma volatility (green line). On the x-axis is the value of the observed quantile, while on the y-axis we report the number of observations.*

in order to validate the observed fluctuations in the degree of multifractality as truly significant, we have shown that the fluctuations observed in empirical data are truly crossing the extreme quantiles expected from a MRW with constant intermittency coefficient. Confidence intervals have been computed by estimating a proxy of multifractality on different realisations of MRW's simulated with parameters obtained from each empirical time series. The analysis performed in this chapter confirms that the tails of the unconditional distribution of stock returns contribute significantly to the observed fluctuations of multifractality, but cannot fully explain the totality of the quantile crossings. Thus the disagreement between model and data could be possibly explained assuming the intermittency parameter to be time varying, which corresponds to having a breakdown in the serial dependence of the volatility. A possible time varying nature of the intermittency

would also affect the tails of the distribution, which depend critically on λ. We have also shown that increasing the intermittency coefficient makes, as expected, the tails of the GHE distribution computed on the synthetic series thicker, while we found no relevant dependence with T. Our analysis thus suggests that a varying intermittency coefficient may be the correct guess towards the inclusion of the observed empirical facts into future multifractal modelling. From an economic point of view, the structure breaks could have various grounds: they could reflect evolving economic fundamentals at different frequencies as well as corporate actions which typically rebounds on the stock prices showing remarkable spree. For example Microsoft Corp share price (exhibiting 33% exceedances against the log-normal MRW benchmark) plummeted sharply from \$178.13 to \$92.38 after a 2 for 1 share split announced in March 1999. Similarly a share split for Citigroup shares (39% exceedances against the log-normal MRW benchmark) in May 2011 caused the price to skyrocket tenfold from \$4.52 to \$44.16. These corporate actions clearly influence the statistical properties of the time series, which possibly affect the multi scaling patterns. An interesting outlook is to look for correlation between economic macro data and multiscaling properties of financial time series, which would reveal to what extent is the financial market reflecting the economic activity of a country.

3.4 Summary

In this chapter we have performed an extensive analysis aimed at validating the dynamical variation of multifractality on a set of NYSE stocks. We have found robust confirmation that, against the framework of the log-normal MRW, a large portion of stocks exhibit time-varying multifractal properties. We have then considered other benchmarks, factoring in most of the effect of the tails, but the hypothesis of constant multifractality is still violated on a relevant percentage of stocks. Overall, the conclusion one can draw from this analysis is that non-stationarity of stock returns rebounds onto their multifractal properties, which are in turn time-varying.

Chapter 4

Interplay between Scaling Properties and Cross-correlation

In this chapter we study the interdependence between multifractal properties of stock returns and their cross-correlation hierarchical properties. We demonstrate that the hierarchical order of stock returns of major international equity markets is remarkably intertwined with their multifractal properties revealing a deep interplay between different aspects of complexity of financial markets. The observation reported in this chapter constitute a new stylized fact [21], presenting a new perspective towards the merging of univariate multi scaling and multivariate cross-correlation properties of financial time series which will hopefully serve as the basis for further unwinding of financial market complexity. As a further study of the relationship between cross-correlation and temporal properties of financial time series, we also investigate the relationship between cross-correlation and volatility autocorrelation and find that stocks filtered through PMFG [99] are also those showing highest autocorrelations, an observation which motivates further the use of the PMFG graph filtering technique to retain relevant information in large datasets.

4.1 Introduction

Financial markets dynamics is driven by *forces* that show their signatures ubiquitously through distinct complex behaviours of the price historical time series. A major challenge is to seek connections between different aspects of complexity and to come up with some common mechanism able to explain these aspects coherently. In particular, we focus on two main elements that define the complexity of financial time series: their multifractality and their cross-correlation. So far, these two manifestation of complexity have been investigated separately. In this chapter we point out that - in fact - they are related. We unveil this hidden interdependence by studying the relationship between degree of multifractality and cross-correlation hierarchical order measured through a clustering algorithm [100, 101].

In the characterisation of financial markets' dependency structure great popularity has been gained over the past fifteen years by network theory approach [102]. The basic idea, introduced in Econophysics by Mantegna [14], is that the dependency properties of a group of assets can be studied after the system has been cast into a network (or a graph), whose links are somehow related to the pairwise cross-correlations. Particularly when the number of assets is very large, network theory tools allow to filter the information enclosed in the correlation matrix according to certain topological constraints. The first (and simplest) constraint, which was introduced in [14], considers the sorted list of all pair cross-correlations and builds a network starting from the most correlated pairs retaining only those pairs that preserve the tree-nature of the graph, so that one ends up with a Minimum Spanning Tree (MST)[1]. The MST is meaningful in providing the backbone of the correlation structure, but has the drawback of not allowing cycles in the graph, which basically rules out any possibility of connecting directly more than two stocks highly correlated between each other. This limitation is overcome by

[1] In graph theory a spanning tree associated with a graph $\mathbb{G}$ is a subgraph that connects all nodes with no cycles. The Minimum Spanning Tree is a spanning tree such that the sum of weights associated to the links is minimised. Since the weights are given by the ultra metric distance $d_{ij} = \sqrt{2(1-\rho_{ij})}$ with ρ_{ij} the correlation coefficient between assets i and j, this corresponds to maximising the sum of correlations. For a graph with N nodes the associated MST has $N-1$ links.

relaxing the constraint on the graph being a spanning tree to it being a planar graph, which allows cycles as long as they don't break planarity of the network. The resulting graph is called Planar Maximally Filtered Graph (PMFG) and was first introduced in [99, 103, 104]. The PMFG is constructed along the same line of the MST with the constraint now being that of preserving planarity, which allows to retain more information [2]. Generalising the embedding of graphs on hyperbolic surfaces in order to retain more information than that preserved by planar graphs is the object of current research - see [105, 106] for the most recent contributions. Further to the network of interaction the filtered correlation matrix of a multivariate set of financial time series can be used to detect a hierarchical structure through clustering algorithms. Clusters give information about the organisation of the system identifying the most correlated assets and grouping them into a hierarchy of clusters and subclusters [107–109]. The detection of clusters of correlated stocks is, in turn, relevant in the identification of distinct risk factors affecting the market heterogeneously [110, 111]. Moreover the correlation based clustering procedure and its corresponding hierarchical tree can be associated to a correlation based network. As an example, it is natural to associate the MST with the Single Linkage Cluster Analysis (SLCA) [112]. In [22] the authors introduce the Deterministic Bubble Hierarchical Tree (DBHT), a new deterministic clustering algorithm starting from the PMFG which is shown to outperform most of existing clustering algorithms in the detection of meaningful clusters. The DBHT provides both the intra-cluster hierarchy, describing the way clusters are composed and the inter-cluster hierarchy, which contains information about the way clusters are organised among them: for details about how the method is constructed we refer to Appendix D and to references [22, 113]. Among several clustering algorithms present in the literature, we have chosen to perform the clustering through DBHT [22].

The chapter is organised as follows: Section 4.2 is devoted to the main empirical result, that is the interdependence between hierarchical order and multifractal

[2] For a graph with N nodes, the total number of links of the PMFG is $3(N-2)$.

properties of stock returns; in the same section we analyse data from all different markets and provide a bootstrapped study on NYSE data which enforce even further the observed behaviour. In Section 4.3 we further show that increasing multifractality in time goes together with increasing hierarchical order and thus the interdependence between the two variables is confirmed also dynamically. Finally, in Section 4.4 we investigate dependency between cross-correlation and volatility autocorrelation.

4.2 Multifractality and Correlation Hierarchical Structure are related

In this section we investigate the interplay between scaling and cross-correlation hierarchical properties of major international equity markets. We looked at different dataset: the first one includes the 342 most capitalised stocks from NYSE in the period 2-01-1997 to 31-12-2012 (this is the same dataset used in Chapter 3); another dataset includes the 185 most capitalised stocks from London Stock Exchange (LSE) in the period 04-01-2000 to 21-08-2013; two further datasets comprise the 386 most capitalised stocks listed in the Tokyo Stock Exchange (TSE) and the 136 most capitalised stocks listed in the Hong Kong Stock Exchange (HKSE), both in the period 01-01-2003 through 26-08-2013. All data have been provided by Bloomberg. The differences in the lengths of the prices time series are due to having retained only the stocks whose prices have been continuously traded in the time period under study, after having used the selection criterium of the 400 most capitalised companies.
Analysis and results presented in this section are part of the paper [21].

4.2.1 NYSE: Sectors and Clusters

We have first performed a DBHT clustering on the NYSE dataset. The clustering is performed on the weighted graph associated with the correlation matrix of

the empirical time series, whose entries are computed via the weighted Pearson estimator defined, for two time series X_i^k and X_j^k, as [83]

$$\hat{\rho}_{ij} = \frac{\sum_{k=1}^{T} w_k (X_i^k - \langle X_i \rangle) \cdot (X_j^k - \langle X_j \rangle)}{\sqrt{\sum_{i=1}^{T} w_k (X_i^k - \langle X_i \rangle)^2 \cdot \sum_{i=1}^{T} w_k (X_j^k - \langle X_j \rangle)^2}} \tag{4.1}$$

where

$$\langle X_i \rangle_w = \frac{1}{T} \sum_{k=1}^{T} w_k X_i^k, \qquad \langle X_j \rangle_w = \frac{1}{T} \sum_{k=1}^{T} w_k X_j^k \tag{4.2}$$

We have implemented, in order to reduce the impact of remote events on present correlations, exponential weights $w_t = w_0 \exp\left(\frac{t-T}{\theta}\right)$ such that $w_t > 0$ and $\sum_{t=1}^{T} w_t = 1$. The parameter θ has been set to $\theta = T/3$ according to criteria previously established in [83], bearing in mind that a θ too little would cast all weight on recent observations, letting correlations loose statistical significance, whereas a too large θ would render the weighting procedure ineffective, as weights would become uniform. For the NYSE data set, where $T = 4026$, the criterium $\theta = T/3$ corresponds to $\theta = 1342$.

The DBHT identifies clusters of stocks and generates a hierarchical organisations both intra-clusters and inter-clusters, which can be visualised by means of a dendrogram. While the intra-cluster hierarchy reveals the dependence between different clusters, the inter-cluster hierarchy provides additional information about the nested organisation of stocks inside each cluster. We will refer to the junctions in the dendrograms as dendrogram nodes (or simply nodes) and to the hierarchical structure above (under) the cluster level as super (sub)-cluster hierarchy respectively.

The DBHT method produces thirteen clusters quite heterogeneous in sizes and compositions. In Table 4.1 we summarise the composition of the clusters in terms of stocks belonging to different sectors. The market sectors we compare the DBHT clusters with are those provided by Bloomberg and are reported in Table 4.2.

Cluster	Size	Financial	Technology	Healthcare	Basic Materials	Industrial Goods	Consumer Goods	Utilities	Services	Conglomerates
1	54		16.7%	16.7%	18.5%	9.3%	**22.2%**		14.8%	1.9%
2	56	**37.5 %**	3.6%			1.8%	14.3%		**41.1%**	1.8%
3	64		6.2%	3.1%		**43.8%**	14%		**25%**	7.8%
4	10	**60%**		30%					10%	
5	20				**75%**	10%		10%	5%	
6	23	**100%**								
7	27		**96.3%**						3.7%	
8	12		**33.3%**		25%		**33.3%**		8.3%	
9	14						**100%**			
10	8				**100%**					
11	10			**80%**					20%	
12	21						4.8%		**95.2%**	
13	23							**100%**		

TABLE 4.1: We report, for each of the 13 clusters identified through the DBHT over the whole time period 2-01-1997 to 31-12-2012, the percentages of the stocks from each of the 9 Bloomberg sectors. The first column to the left assigns labels to each of the clusters. On the second column to the left we also show the size of each cluster, that is the number of stocks each cluster contains.

Sector label	Sector	Number of stocks
1	Basic Materials	36
2	Conglomerates	7
3	Consumer Goods	48
4	Financial	50
5	Healthcare	22
6	Industrial Goods	36
7	Services	73
8	Technology	45
9	Utilities	25

TABLE 4.2: **Number of stocks in each sector.**

We find three large clusters (labelled 1,2 and 3) containing $54, 56$ and 64 stocks respectively and all quite heterogeneous in composition. Clusters labelled by 2 and 3 show a relevant fraction of stocks belonging to Financial and Industrial sectors. This is confirmed by the enrichment analysis [114, 115], which validates the identification of the clusters with the respective sectors at a confidence level of 1%. The enrichment test allows to assess whether a cluster is identifiable with a sector at a certain confidence level. The other clusters are significantly smaller, but still capture quite well the sector membership: cluster 13 for example contains stocks from the Utility sector only, cluster 9 and 10 from Consumer Goods and Basic Materials sectors respectively. Cluster 7 has a very large component of stocks belonging to the Technology sector, while cluster 12 is mainly composed by stocks from the Services.

The DBHT method was singled out among the many clustering algorithms available after having verified that it works better than other methods including the Single Linkage Cluster Analysis (SLCA) [112] in recovering a well-diversified hierarchical structure as well as in identifying clusters uniquely. SLCA in fact does not distinguish between clusters of different order, while DBHT distinguished clusters from, for instance, sub-clusters. Further to this, we verified that, on our datasets, SLCA tends to produce very large hubs of stocks that bias the correlation hierarchical order, a concept which is introduced in the following section.

4.2.2 Hierarchical Paths and Trees

The hierarchical structure produced by the DBHT provides a natural way to associate with each stock a hierarchical order, that is the number of nodes above each stock along the path from the stock at the bottom of the tree to the top. We will denote with n the generic order and with n_i the hierarchical order of a specific stock i, where $i = 1, \ldots, N$. In other words, n_i measures how deep down the hierarchical tree lies a certain stock i. A schematic hierarchical structure is given in

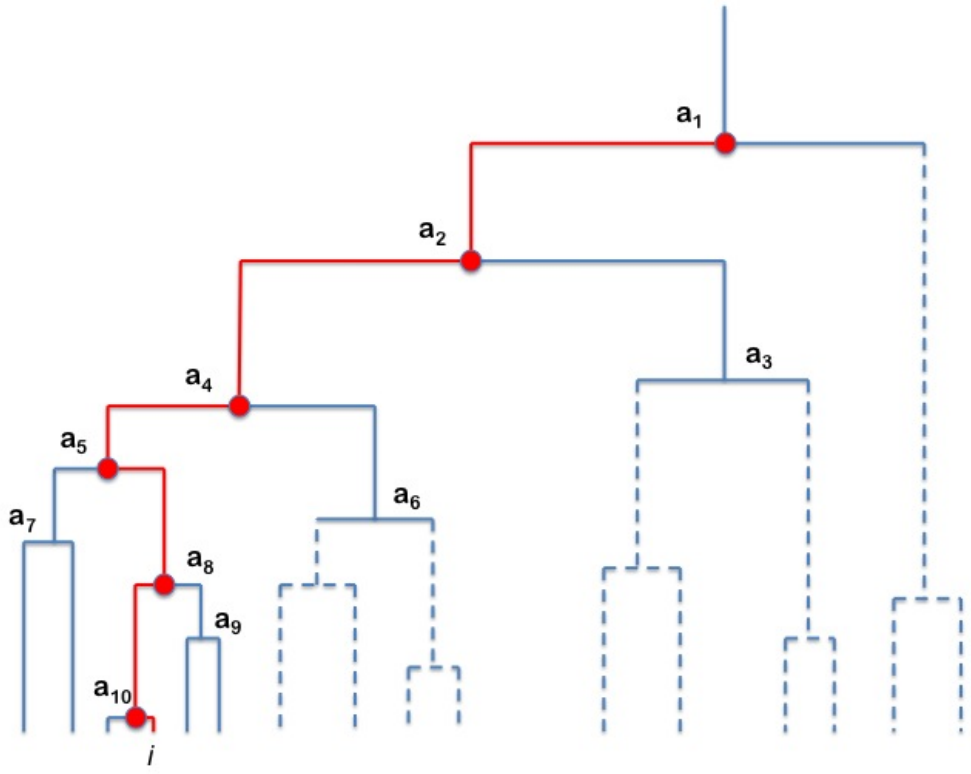

FIGURE 4.1: **Example of hierarchical structure.** The path highlighted in red is Γ_i, while the thick red bullets are the nodes $a_m,\ m \in \gamma_i$, which correspond to the risks stock i is exposed to. Note that the risk a_1 is common to all stocks, while other risks affect groups of stocks only. The dashed branches of the dendrogram indicate an arbitrary hierarchy.

Figure 4.1. The hierarchical tree Γ_i is defined to be the set of nodes along the simple path γ_i from the stock to the top of the dendrogram, i.e. $\Gamma_i = \{a_m : m \in \gamma_i\}$, where we denote by γ_i the set of positive numbers identifying the nodes above stock i. Note that there's only one such simple path for each stock i. Then the hierarchical order is defined as the cardinality of Γ_i, $n_i = \text{card}(\Gamma_i)$. In the example

reproduced in Figure 4.1, for the stock labeled by i, we have $\gamma_i = \{1, 2, 4, 5, 8, 10\}$, $\Gamma_i = \{a_1, a_2, a_4, a_5, a_8, a_{10}\}$, i.e. the set of nodes denoted by the red dots and $n_i = 6$.

4.2.3 NYSE: Results

To assess the level of entanglement between multifractal properties and cross-correlation hierarchical structure, we looked at how the hierarchical order n behaves with respect to $\Delta H(1,2)$, the proxy of multifractality which was defined in Chapter 2. We have removed from the analysis all stocks whose multifractality cannot be statistically distinguished from zero, which corresponds to weak multifractal behaviour and hence would not be relevant in this context. Specifically, we have considered as significantly multifractal only those stocks showing $\Delta H(1,2) > 0.015$. This value has been chosen as the 95% probability value of $\Delta H(1,2)$ on uniscaling processes. To obtain it we have simulated 1000 realisations of fractional Brownian motion[3] with Hurst parameter H randomly chosen in the range $[0.1, 0.9]$ and then computed the value of $\Delta H(1,2)$ corresponding to (two side) 95% probability. In our NYSE dataset we found 72 stocks excluded as weakly multifractal.

After having selected the relevant stocks, we first looked at the average behaviour observed in the market. We have computed the average values of $\Delta H(1,2)$ for each different order $\bar{n}$ as

$$\langle \Delta H(1,2) \rangle_{\bar{n}} = \frac{1}{N_{\bar{n}}} \sum_{i:n_i=\bar{n}} \Delta H_i(1,2), \tag{4.3}$$

where $N_{\bar{n}}$ is the number of stocks with order $\bar{n}$. Likewise the standard deviation has been computed as

$$s_{\bar{n}} = \sqrt{\frac{1}{N_{\bar{n}} - 1} \sum_{i:n_i=\bar{n}} (\Delta H_i(1,2) - \langle \Delta H(1,2) \rangle_{\bar{n}})^2}. \tag{4.4}$$

[3] A notoriously uniscaling process.

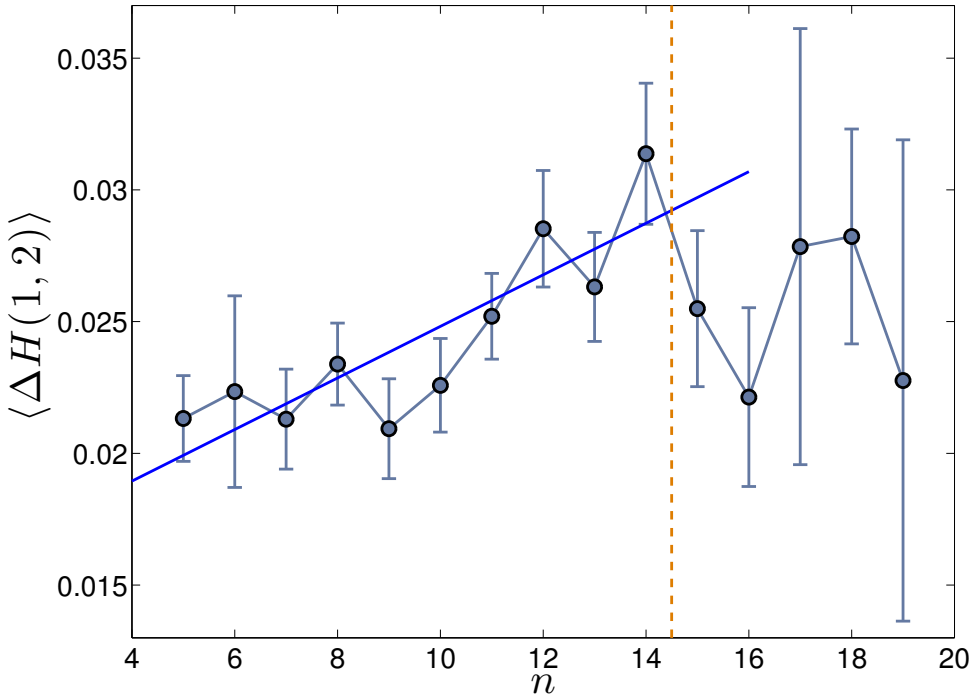

FIGURE 4.2: **Demonstration that multifractality and hierarchical order are positively correlated.** (Color online) We plot in blue circles with error bars the multifractal indicator $\langle \Delta H(1,2) \rangle$ averaged over the stocks sharing the same hierarchical order, against the hierarchical order n. The blue solid line is the linear fit over the averages, while the orange horizontal dashed line marks the limit up to which the increasing trend is observed. The error bars are the standard errors computed as $s/\sqrt{N}$, where s is the standard deviation over the stocks having same hierarchical order. Correlation coefficient between the two variables and corresponding p-value are (0.8,0.002).

We plot in Figure 4.2 the mean value $\langle \Delta H(1,2) \rangle$ with standard error $s/\sqrt{N}$ (where s is the standard deviation on the mean) for each observed hierarchical order on all stocks analysed. We observe a positive dependence between the two variables up to $n = 14$, followed by some noisier flat trend. The positive correlation between $\langle \Delta H(1,2) \rangle$ and n is 0.8 and its significance is also confirmed by performing a t-test that has returned p-value $p = 0.002$. All stocks with hierarchical order in the range $[5, 14]$ (which accounts for 90% of all stocks) exhibit multifractal properties increasing along with their depth in the hierarchy of correlations. On the other hand, the apparent saturation observed for orders larger than 14 suggests that the hierarchical structure of cross-correlations may be responsible for the multifractal properties of the stocks only up to a certain order. Let us note in particular that

the number of stocks found with hierarchical order $n > 14$ is too small to allow any robust statistical conclusion, which is also the reason why standard errors in Figure 4.2 are very large for $n > 14$.

As a further step, we have looked for the same positive dependence between hierarchical order and multifractality on specific market sectors and on clusters found through DBHT clustering algorithm. Operatively, the averages in Equations (4.3) and (4.4) over the entire dataset are replaced by averages over stocks belonging to specific sectors and clusters. We show in the top of Figure 4.3 plots of the multifractality indicator versus the hierarchical order computed on stocks belonging to Financial and Industrial sectors. The black dots are values for single stocks whereas the red squares are the average multifractal indicators for each order. We also plot the best fit on the dots (thick blue line). Both sets of stocks show a very well defined positive correlation between the hierarchical order n and the multifractality indicator $\Delta H(1, 2)$, which is evident from the positive trend recovered in both examples. Again, the dependence is found to be significant with p-value $p = 0.02$.

In Figure 4.3 we also report the trends observed on two of the largest clusters found through the DBHT. The same positive correlation between multifractality and hierarchical order is observed on the clusters best identifiable with the corresponding sectors: the Financial sector has large components in clusters $2, 4$ and 6, whereas a large component of stocks from the Industrial sector is found in cluster 3 (see Table 4.1). All stocks belonging to the largest clusters and the corresponding sectors that better match the clusters show a tendency to pair high multifractality with depth in the hierarchy of correlations. The positive correlation between multifractality and n has been validated via the t-test already performed for the entire set of stocks shown in Figure 4.2. For the plots in Figure 4.3 we found for correlation and p-value (clockwise from top left plot): (0.48,0.0007), (0.61,0.05), (0.30,0.03) and (0.24,0.11). These results confirm the statistical significance of the trends observed, apart from the case of cluster no. 3. Note that, however, although in all cases shown multifractality is found to depend on the hierarchical order, this dependence is likely to be non-linear. Let us consider for example the

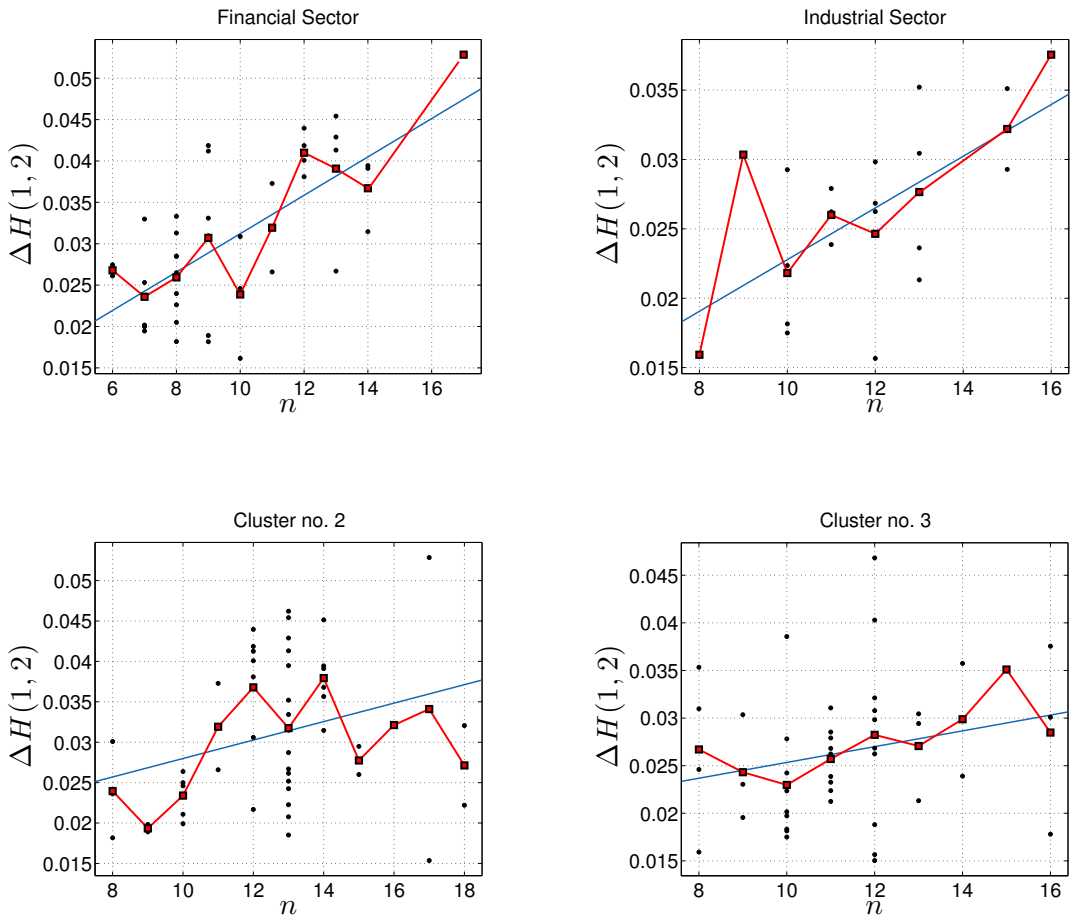

FIGURE 4.3: We plot in black dots $\Delta H(1,2)$ against the hierarchical order n for stocks in the Financial (top left) and Industrial (top right) sectors and cluster 2 (bottom left) and cluster 3 (bottom right). In all plots the blue line is the best fit of the dots, while the red squares are the averages of $\Delta H(1,2)$ for each fixed order. Correlation coefficients and p-values are respectively (clockwise from toppleft plot): (0.48,0.0007), (0.61,0.05), (0.30,0.03) and (0.24,0.11).

stocks in cluster 2: the trend is clearly increasing for $n \in [9,14]$, then there is some oscillation around somewhat less than the value of multifractality attained for $n = 14$ similarly to what observed on the complete set of stocks in Figure 4.2. A possible way to unveil these non-linearities in a more evident fashion is by restricting the attention on sub-clusters, i.e. groups of stocks below the cluster level in the hierarchy. Restricting to hierarchical levels below the clusters has two opposite consequences: on the one hand it allows to be more strict in factoring out a larger number of common hierarchical levels, hence making the differentiation between different stocks more relevant. On the contrary, it can reduce significantly the number of stocks and thus potentially increase the influence of noise on the

trends. After a thorough analysis of all sub-clusters of different order, we found that in all cases the trends recovered on sub-clusters are, when significant, the same observed on the clusters. We use the notation $\mathcal{S}^j_{i,k}$ to label such sets, where $j = 1, \ldots, 13$ denotes the cluster label, i denotes the order below the cluster and $k = 1, \ldots, 2^i$ denotes which of the 2^i sub-clusters at each order we are considering. Plots for the behaviour $\Delta H(1,2)$ *vs* n are reported in Appendix F. We plot in Figure E.1 in Appendix F $\Delta H(1,2)$ versus n for cluster no.6 in Table 4.1 and for two sub-clusters $\mathcal{S}^3_{1,1}$ and $\mathcal{S}^2_{2,1}$. The sub-cluster $\mathcal{S}^3_{1,1}$ is very similar to cluster no.3 shown in Figure 4.3: cluster no.3 in fact results from the merging of $\mathcal{S}^3_{1,1}$ and $\mathcal{S}^3_{1,2}$, the latter one containing only 5 stocks.

4.2.4 Bootstrapping the DBHT

In this section we study the robustness of the result presented in Section 4.2.3. Specifically, we study the robustness of the DBHT method by validating the dendrogram obtained through a bootstrapping procedure [116, 117]. The estimates whose accuracy need to be validated are the hierarchical orders of the stocks. Even slight modifications in the ranking of correlations in fact are likely to undermine the organisation recovered with the original configuration and to return hierarchical orders very far from those observed on real data. We proceed as follows: we construct a number of synthetic resamples of the original dataset, where each time the surrogate data are obtained by the usual procedure of sampling with replacement. Each time the statistics of interest is computed and at the end the original observation is validated against the distribution of the resample estimates obtained from the surrogate series. Since we are interested in giving accurate estimates of the orders of all stocks, the DBHT is performed on each bootstrapped correlation matrix for a total of 1000 resamples. From each of these dendrograms constructed from surrogate correlation matrices we extract the vector of bootstrapped hierarchical orders $\mathbf{n}_\beta = (n^1, n^2, \ldots, n^N)_\beta$, where $N = 342$ is the number of stocks in the dataset and $\beta = 1, \ldots, 1000$. Then we compare the vector of

hierarchical orders observed on the empirical time series with the synthetic distribution of bootstrapped values. For each stock, we discard the empirical order as non-reliable if it falls beyond the two-sided 0.9 level of probability obtained from the cumulative distribution function of the bootstrapped series. As an example of the validation procedure, we report the case of the order of the stock Air Products and Chemicals INC in Figure 4.4. We plot the cdf of the bootstrapped hierarchical orders and the empirical order n in pink vertical line. The hierarchical order measured on the original DBHT is not validated by the bootstrap test as it falls beyond the 0.9 probability level.

The number of stocks which fail the bootstrap test is remarkably high and, by coincidence, is exactly half of the total number of stocks. We show in Figure 4.5 the distribution of the measured order before the bootstrapping (blue bars) and the distribution of the order of the stocks validated (dashed red line). Although the number of stocks is halved, the shape of the distribution remains similar. We then denote the set of valid hierarchical orders as n_B.

After the bootstrap validation, the trends recovered in the plots of $\Delta H(1,2)$ versus n reveal some further hidden structure. We show in Figure 4.6 the trends observed on clusters 2, 3 and sub cluster $\mathcal{S}_{1,1}^{3,B}$, where the stocks shown are only those that have been validated via bootstrapping. As also shown in Figure 4.5, the stocks excluded are more or less uniform in their order, which results in stocks with very large (above 15) and very small (below 7) order to almost disappear from the set. Although the number of stocks is reduced, the dependence of multifractality on the hierarchical order is seen even more neatly. Moreover the non-linear behaviour guessed from the plots presented in the previous section is now even more evident from the middle plot in Figure 4.6.

4.2.5 London Stock Exchange

The previous results found on NYSE data have been verified and confirmed on the LSE dataset. We have analysed the set of 185 most capitalised stocks traded at the London Stock Exchange (LSE) in the period 04/01/2000-21/08/2013. The

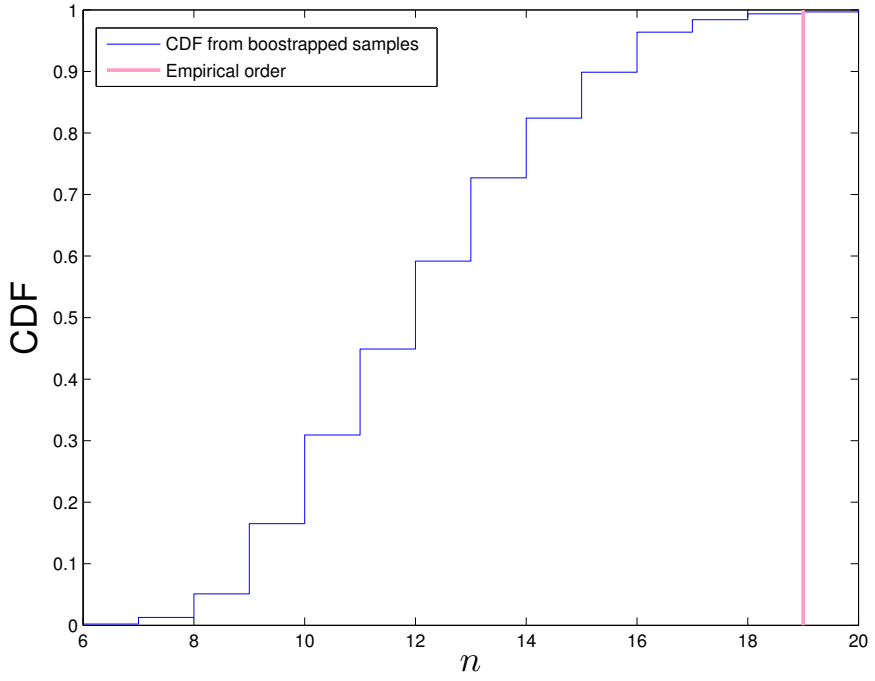

FIGURE 4.4: We plot (blue line) the cumulative distribution function of the bootstrapped hierarchical order for the stock Air Products and Chemicals INC. The pink vertical line corresponds to the hierarchical order of the same stock obtained from the DBHT on the original dataset. Its level of probability is such that the estimated order cannot be validated.

DBHT returned a range of hierarchical order comparable to that obtained for the NYSE data, with stocks exhibiting $n \in [4, 17]$. As done for NYSE, we have looked at the behaviour $\langle \Delta H(1,2) \rangle$ *vs* n, where the average $\langle \Delta H(1,2) \rangle$ is taken over all stocks exhibiting the same hierarchical order. We plot in Figure 4.7 $\langle \Delta H(1,2) \rangle$ as a function of the hierarchical order n. We observed the same increasing trend up to $n = 12$, followed by a drop and a plateau at higher orders, where the observations are however fewer. The regime $n \in [4, 12]$ confirms the observation reported in the main text for NYSE: within a certain range of orders, multifractality and hierarchical order appear to be positively correlated. Even more than in the NYSE example though, the dependency between the two quantities shows non-linear features.

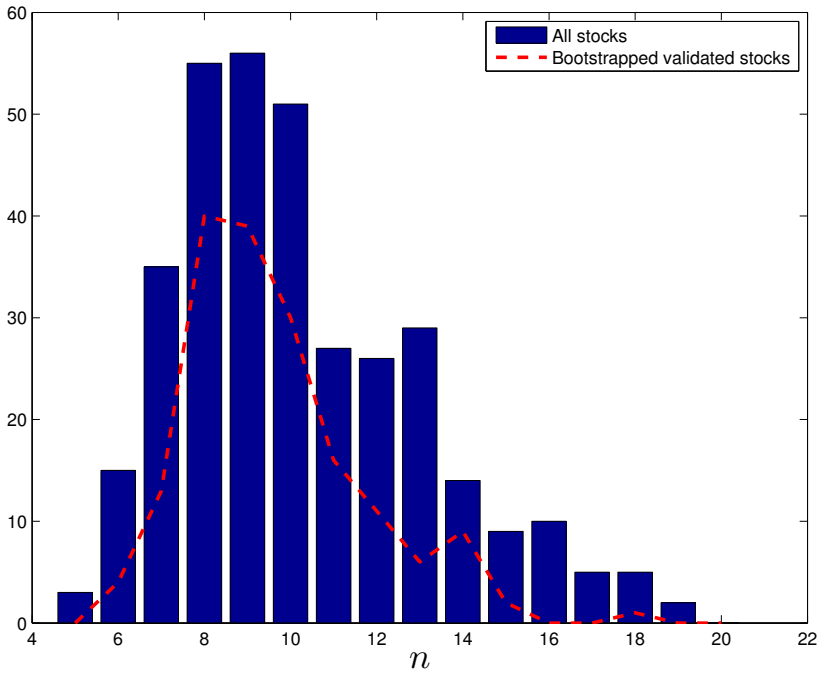

FIGURE 4.5: We plot in blue bars the histogram of the hierarchical orders obtained from the DBHT for the complete dataset. The thick dashed red line is the histogram of the hierarchical orders for those stocks that passed the bootstrapping validation.

4.2.6 Asian Markets: one dominant market mode

We have also analysed the set of 386 most capitalised stocks listed in the Tokyo Stock Exchange (TSE) and the 136 most capitalised stocks in the Hong Kong Stock Exchange (HKSE) in the period 01/01/2003-26/08/2013. We report in Table 4.3 the most relevant features observed on the two Asian datasets: number of clusters, average correlation coefficient and hierarchical order range. For both markets, the DBHT returns a taxonomy of clusters very different from that found in the NYSE and LSE data. In TSE data we found only six clusters compared to the thirteen of NYSE and ten of LSE. Among these clusters there is a gigantic cluster containing 300 stocks, while the remaining stocks are more or less uniformly distributed in the remaining five clusters. A closer inspection of the correlation matrix of this system reveals that, on average, these stocks show correlation larger than those observed in the western markets, with an average correlation of 0.37 (compared to 0.18 found in NYSE and 0.20 for LSE). The average correlation observed on

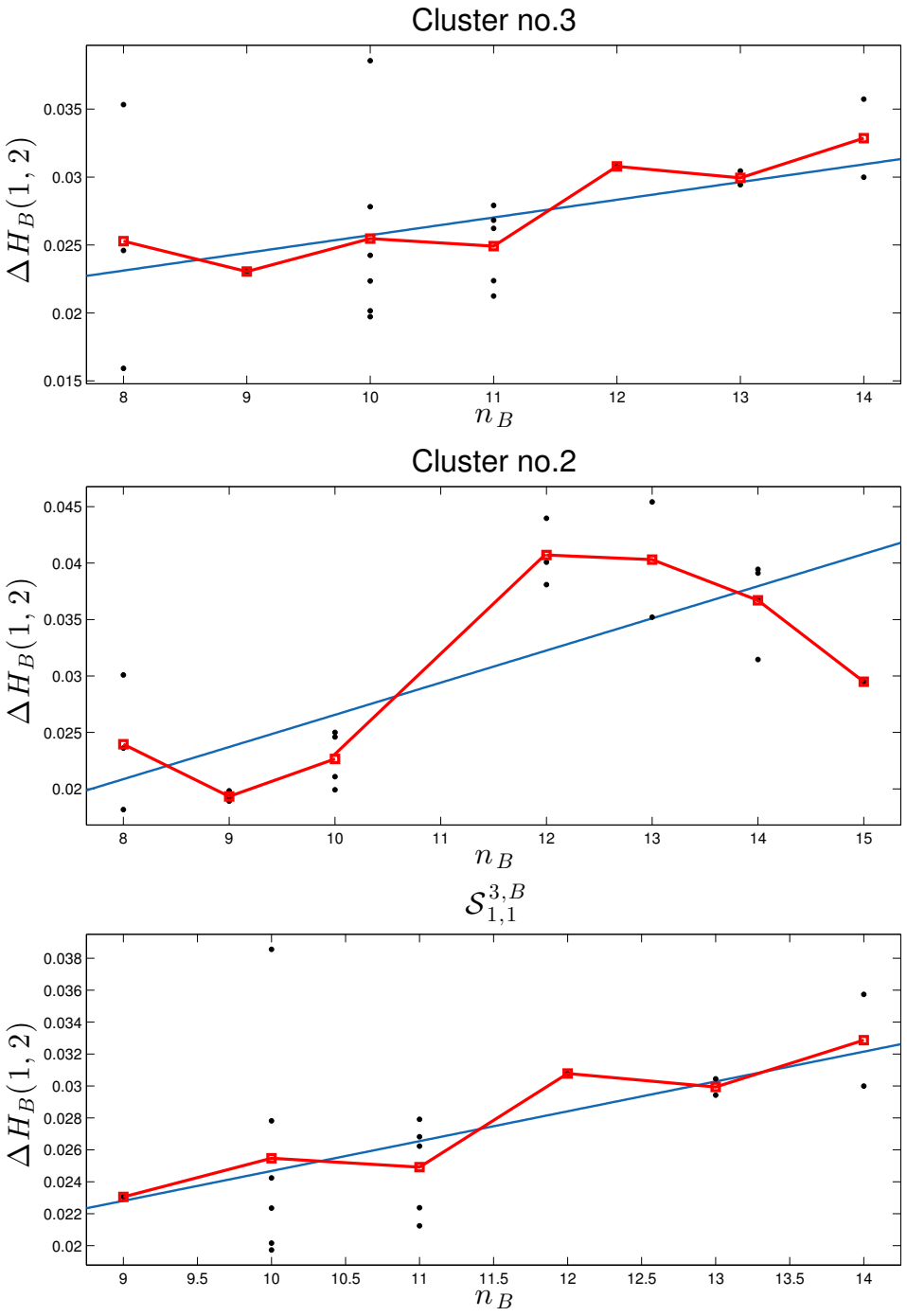

FIGURE 4.6: $\Delta H(1,2)$ against the hierarchical order n for stocks in cluster no.3 (top), cluster no.2 (middle) and sub cluster $\mathcal{S}_{1,1}^{3,B}$ (bottom) for stocks whose order has been validated via bootstrapping. In all three plots the blue line is the best fit of the dots, while the red squares are the averages of $\Delta H(1,2)$ for each fixed order. Correlation coefficients and corresponding p-values are (top to bottom): (0.84,0.02), (0.67,0.06) and (0.85,0.02).

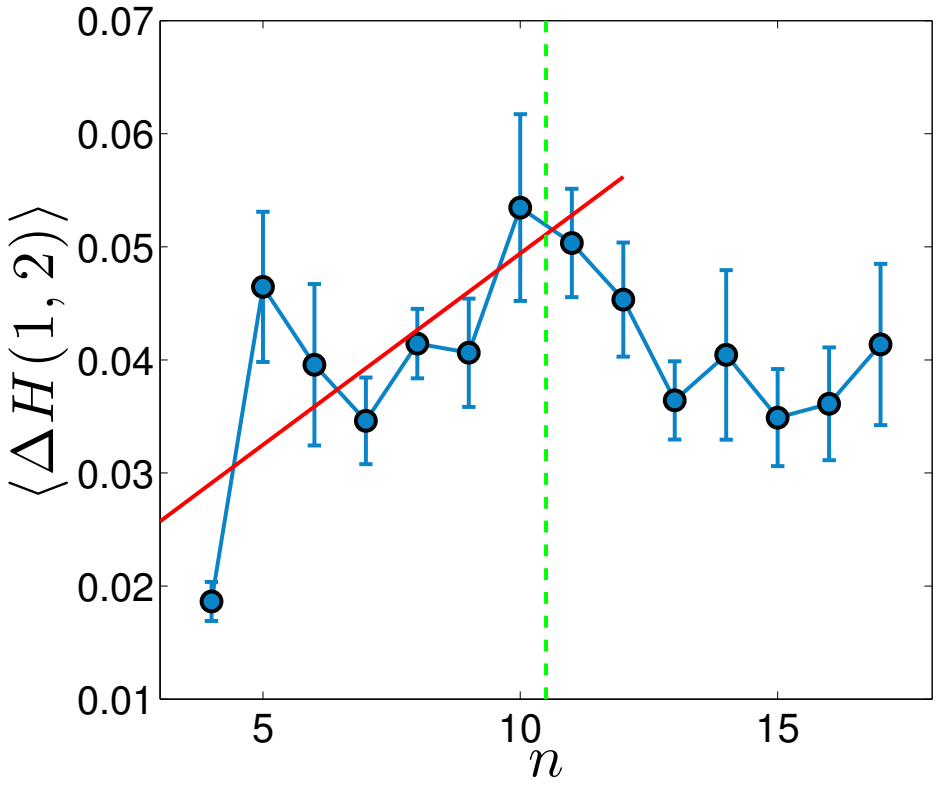

FIGURE 4.7: Positive correlation between multifractality and hierarchical order for stocks traded in LSE. We plot in blue circles the average $\langle \Delta H(1,2) \rangle$ for each observed hierarchical order against the corresponding n. Error bars are standard errors on the mean, computed as $s/\sqrt{N_n}$, with N_n the number of stocks having order n and s the standard deviation on the mean. The red line is the best linear fit on the averages and the vertical dashed green line marks the end of the linear trend.

HKSE data is also sensibly larger than that observed on western markets. As a result of this very homogeneous taxonomy of clusters, stocks exhibit hierarchical order varying in the range $[3, 121]$: a dendrogram with a small number of clusters necessary entails the hierarchy to develop vertically. For the HKSE we found six clusters and $n \in [4, 51]$. It is nearly impossible to distinguish the effect of such a wide range of orders on the multifractal properties of stocks, but nonetheless the interdependence between the $\langle \Delta H(1,2) \rangle$ and n is still observed for small n: we show in Figure 4.8 $\langle \Delta H(1,2) \rangle$ as a function of n for TSE (top) and for HKSE (bottom). In both datasets the two quantities show some form of positive dependence. We mention that the number of observations for each n is very small to allow any robust statistical conclusion (the errors are consequently very large),

	TSE	HKSE
Number of Clusters	6	6
Average Correlation	0.37	0.29
Order Range	[3,121]	[4,51]

TABLE 4.3: We show in this table the number of clusters, average correlation and hierarchical order range for the two Asian data sets.

but the effect of the stacked risks seems to be visible also in this case.

To remove the presence of the very large hubs of correlated stocks, we have performed a detrending of the time series in order to get rid of the dominant market mode and thus observe a more heterogeneous structure of clusters. Returns are considered to follow a one factor linear model of the following form [27, 118]

$$r_{i,t} = a_i + \beta_i I_t + c_{i,t} \tag{4.5}$$

where $I_t = 1/N \sum_{i=1}^{N} r_{i,t}$ is the composite index with homogeneous weights. After estimating the coefficients β_i for each stock, the correlation matrix has been computed on the residuals $c_{i,t}$, thus deprived of any common trend.

We detrended the TSE data through Equation (4.5) and, as expected, the number of clusters increased from 6 to 21 and the range of n was found to be $[5, 21]$. We plot $\langle \Delta H(1,2) \rangle$ as a function of n for the detrended TSE data in Figure 4.9. One can see that multifractality appears to be completely uncorrelated with n, the behaviour being overall flat and revealing non particular trend. This observation has in our opinion the following interpretation: in a very correlated data set, the market mode is definitely the most relevant factor affecting the dynamics of the stocks. As a consequence, removing the common trend almost completely removes the most relevant source of complexity of the time series, whose remaining heterogenous properties cannot explain the correlation between hierarchical order and multifractality. The Japanese market thus appears to be much less diversified than the American and British ones and one factor may be enough to explain its dynamics.

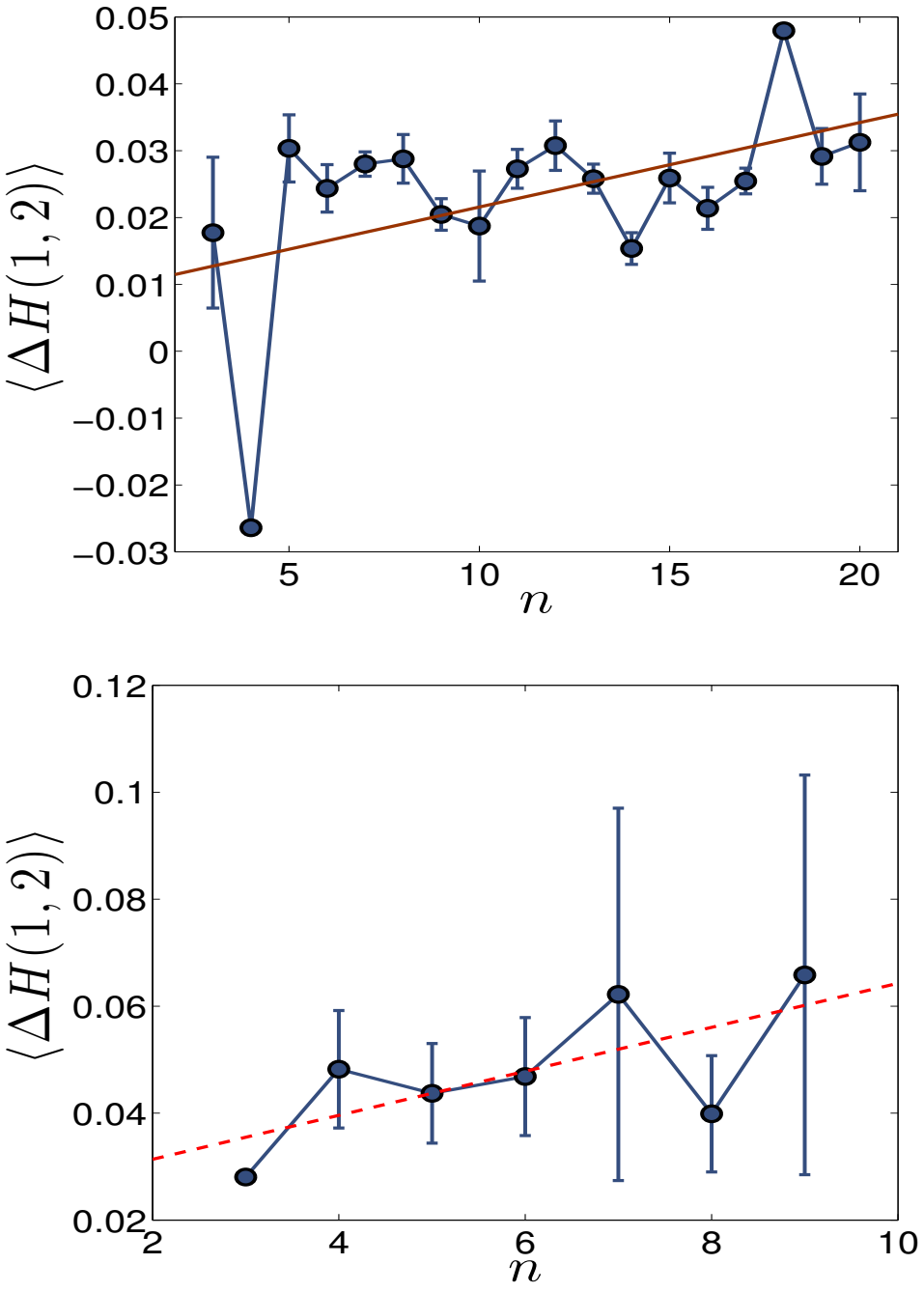

FIGURE 4.8: (Top) Relationship between $\langle \Delta H(1,2) \rangle$ and n for $n \in [3,20]$, neglecting all higher hierarchical orders ($n \in [21,121]$) for TSE. (Bottom) Relationship between $\langle \Delta H(1,2) \rangle$ and n for $n \in [3,9]$, neglecting all higher hierarchical orders ($n \in [10,51]$) for HKSE. The solid lines are the best fits on the averages. In both cases a positive trend is observed. Correlation coefficients and p-values are: (0.41,0.09) for TSE and (0.61,0.14) for HKSE.

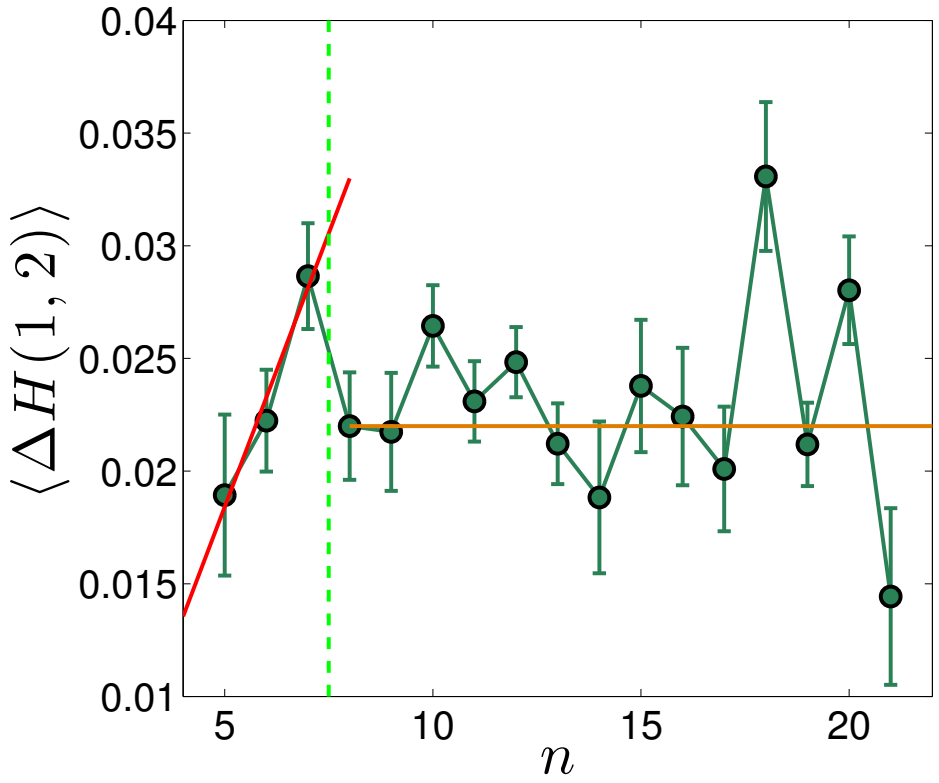

FIGURE 4.9: **Correlation between multifractality and hierarchical order after detrending for stocks traded in TSE.** We plot in green circles the average $\langle \Delta H(1,2) \rangle$ for each observed hierarchical order against the corresponding n. Error bars are standard errors on the mean, computed as $s/\sqrt{N_n}$, with N_n the number of stocks having order n and s the standard deviation on the mean. The red line is the best linear fit on the averages for $n \in [5,7]$, while the orange line is the fit for $n \in [8,21]$. The dashed line marks the end of the initial increasing trend.

As a general comment large correlation hubs thwart the possibility of tracking the effect of the hierarchy on the multi-scaling properties of the time series. In markets diversified enough though, like the NYSE, the effect of many risk factors is still reasonably well detected by the DBHT.

4.3 Dynamical Dependence and Market Coalescence over the Crisis

The correlation between multifractality and hierarchical order has been also detected dynamically in time on NYSE data. We have performed DBHT clustering in time on 50 overlapping time windows of length 752 days. The size of the time window, corresponding to three years of trading time, has been chosen in order to allow for a robust enough statistics, at the same time diversifying as much as possible among different correlation regimes.

We tracked the number of clusters $\mathcal{N}_{c,t}$ observed on the time window t. This gives a measure of how compact the hierarchy is: a large number of clusters corresponds to a hierarchy sprawled horizontally, while a small number of clusters corresponds to a hierarchical tree narrow and deep and therefore compact. We plot in the left plot in Figure 4.10 the observed behaviour of $\mathcal{N}_{c,t}$: each blue circle corresponds to the number of clusters $\mathcal{N}_{c,t}$ resulting from performing the DBHT on the time window t. The hierarchical structure evolves significantly in time revealing time varying topological properties of the dendrograms, as also documented in detail in [115]. We observed a systematic decrease of $\mathcal{N}_{c,t}$ with time, with the trend being steeper in the period preceding the 2007-2008 financial crisis. Hence, the overall contraction of the market already reported in other studies [119–121] corresponds to a contraction of the hierarchy, which reflects the fact that the market tends to show less heterogeneity in correlations. Note that to the shrinkage of the hierarchical structure corresponds an increase of the average correlation $\langle\rho\rangle_t$ in the market, as shown in the right plot in Figure 4.10, where we also report in dashed lines the quintiles of the empirical distribution of correlation coefficients. Remarkably, the increasing coalescence of the hierarchical structure is followed through by an increase in both average multifractality $\langle\Delta H(1,2)\rangle_t$ and average hierarchical order $\langle n\rangle_t$, whose behaviours in time are reported in Figure 4.11. We plot on the left of Figure 4.11 the dynamical evolution of $\langle\Delta H(1,2)\rangle_t$ (where the average $\langle\cdot\rangle$ now is over all stocks) and on the right the dynamical evolution of the average hierarchical order $\langle n\rangle_t$. Note also the sudden synchronous drop in both $\langle\Delta H(1,2)\rangle_t$ and

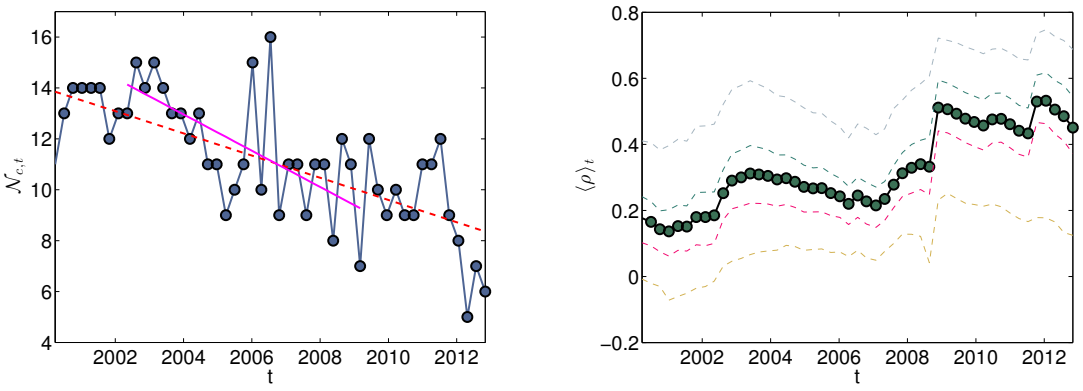

FIGURE 4.10: (Left) The number of cluster $\mathcal{N}_{c,t}$ as a function of time is plotted in time in blue empty circles. The dashed red line is the best fit over the entire time period 2-01-1997 to 31-12-2012, while the magenta line is the fit over the shorter period preceding the 2007/2008 financial crisis September 2002 through November 2007. Standard errors on the circles are not reported as too small to be visible. (Right) Dynamical evolution of the average correlation (thick dots) in the same time period. The dashed coloured lines represent the $2.5\%, 25\%, 75\%, 97.5\%$-quantiles, taken from the distribution of all the observed correlation coefficients.

$\langle n \rangle_t$ around 2008. The increase in multifractality can be explained through the exposure of the prices to a larger set of risks, which seem to be well captured by the increase in the average hierarchical order.

This analysis reveals that the overall increase in complexity of the market shows itself in different forms: increase in multifractality, increase in correlations and compaction of the hierarchical structure, which corresponds to the appearance of higher hierarchical orders. All these interlinked phenomena regard distinct properties of stock returns, yet they seem to remarkably go together. In Chapter 6 we will introduce a model able to account for all these features.

4.4 Dependence between Cross-correlation and Autocorrelation

As a further study of the interplay between cross-correalation and temporal dependence properties of time series, we investigate in this section the relationship between correlation across stocks and the square returns autocorrelation $C_{r^2}(\ell)$.

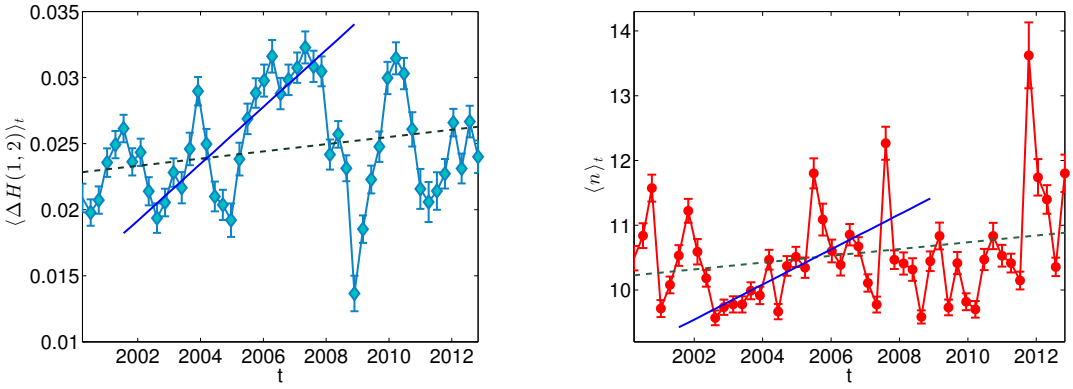

FIGURE 4.11: (Left) The blue diamonds are values of the average multifractality over 50 overlapping time windows. The dashed green line is the best fit over the entire time period 2-01-1997 to 31-12-2012, while the blue line is the fit over the shorter period preceding the 2007/2008 financial crisis, September 2002 through November 2007. (Right) The red circles are values of the average hierarchical order over 50 overlapping time windows. The dashed green line is the best fit over the entire time period 2-01-1997 to 31-12-2012, while the blue line is the fit over the shorter period preceding the 2007/2008 financial crisis, September 2002 through November 2007. In both plots the error bars are the standard mean error on the mean $s/\sqrt{N}$, with s the standard deviation.

Our purpose is that of ascertaining whether highly correlated stocks are also those showing long-ranged autocorrelation in the proxy of volatility r^2. This question was recently tackled in [122], where the author showed that the average cross-correlation is correlated with the volatility autocorrelation decay exponent. Here we use a different approach, studying how $C_{r^2}(\ell)$ correlates with several measures of cross-correlation for each stock i, as function of the time lag ℓ. We find confirmation of the results presented in [122] through a different analysis which also highlights the relevance of planar graphs filtering techniques in identifying the most relevant components of the market.

We consider average values or "global" measures of correlation for each particular stock, which makes it possible to rank each stock among all other stocks under analysis. This is because, whereas autocorrelation is a property of each stock, cross-correlation coefficient involves two stocks and therefore comparing the two quantities is altogether ambiguous. We are not interested in random high correlation between stocks, but rather we would like to distinguish those stocks showing

higher correlation with many other stocks from those other stocks being less correlated. We propose several possible variables to look at.

4.4.1 Definition of the Variables

The first possibility, also used in [122], is the average correlation of a stock i with all other stocks, namely

$$\bar{\rho}_i = \frac{1}{N-1} \sum_{j=1}^{N-1} \hat{\rho}_{ij} \qquad i = 1, \dots, N, \tag{4.6}$$

where $\hat{\rho}_{ij}$ is the measured correlation between stocks i and j, evaluated via the Pearson's estimator. This variable measures how correlated on average the selected stock i is with all the others and it is therefore a global measure of correlation for i.

Another possibility is to look at the number of correlation coefficient being larger than a certain threshold $\tilde{\rho}$. This means looking, for each stocks, at how many correlations are big enough to be considered significant. Formally we can write

$$\mathcal{N}_i(\tilde{\rho}) = \frac{1}{N-1} \sum_{j=1}^{N-1} \mathbb{1}_{\{\hat{\rho}_{ij} > \tilde{\rho}\}}, \qquad i = 1, \dots, N. \tag{4.7}$$

This variable depends on how we choose the threshold. For a high threshold, say from $\tilde{\rho} \sim 0.5$ upwards, many stocks would exhibit $\mathcal{N}_i(\tilde{\rho}) = 0$, hence limiting the set of stocks which could be analysed. On the other hand, if we chose a less strict value, we would retrieve most stocks showing $\mathcal{N}_i(\tilde{\rho}) \neq 0$, but the threshold would be not as strict as to guarantee that these stocks are more correlated than the average.

A third possible interesting measure is to look at those correlation obtained through graph filtering techniques. In particular, we looked at planar maximally filtered graphs ($PMFG$). We considered as a measure of global correlation the average of

the correlation $\hat{\rho}_{ij}^{PMFG}$ survived to the $PMFG$ filtering, that is

$$\bar{\rho}_i^{PMFG} = \frac{1}{n_i^{PMFG}} \sum_{j=1}^{n_i^{PMFG}} \hat{\rho}_{ij}^{PMFG}, \qquad i = 1, \dots, N \tag{4.8}$$

where n_i^{PMFG} is the number of correlation coefficients between stock i and any other stock, filtered via the $PMFG$ procedure.

The three global correlation measures can be written in a more compact form as

$$\mathcal{V}_\alpha = \begin{cases} \bar{\rho}, & \text{if } \alpha = 1 \\ \mathcal{N}(\rho_{ij}; \tilde{\rho}), & \text{if } \alpha = 2 \\ \bar{\rho}^{PMFG}, & \text{if } \alpha = 3. \end{cases} \tag{4.9}$$

With this definition we can define, for each fixed lag ℓ, the correlations between each of the three variables and $C_{r^2}(\ell)$ as

$$\mathbf{C}_\alpha = \mathrm{Corr}(\mathcal{V}_\alpha, C_{r^2}(\ell)) = \begin{cases} \mathrm{Corr}(\bar{\rho}, C_{r^2}(\ell)), & \text{if } \alpha = 1 \\ \mathrm{Corr}(\mathcal{N}(\rho_{ij}; \tilde{\rho}), C_{r^2}(\ell)), & \text{if } \alpha = 2 \\ \mathrm{Corr}(\bar{\rho}^{PMFG}, C_{r^2}(\ell)), & \text{if } \alpha = 3. \end{cases} \tag{4.10}$$

$\mathbf{C}_\alpha$ gives a measure of how the autocorrelation function for a given lag ℓ is entangled with the cross-correlation measure $\mathcal{V}_\alpha$.

4.4.2 Results

We investigated the behaviour of $\mathcal{V}_\alpha$ *vs* $C_{r^2}(\ell)$ for $\alpha = 1, 2, 3$ on the NYSE dataset. We plot in Figure 4.12 $\bar{\rho}$ versus $C_{r^2}(\ell)$ for $\ell = 2, 10, 15, 20, 25, 30$. One can see how the cloud of dots corresponding to each stock seems to describe a non-linear increasing trend, saturating somewhere around $C_{r^2}(\ell) \sim 0.2$, at least for $\ell > 10$. In Figure 4.13 we plot $\mathcal{N}_i(\tilde{\rho} = 0.4)$ *vs* $C_{r^2}(\ell)$ for $\ell = 2, 10, 15, 20, 25, 30$. Here we spot no evident dependence between the two quantities, with the cloud of dots being much scattered. We also looked at other values of the threshold $\tilde{\rho}_i$ - specifically $\tilde{\rho}_i = 0.25, 0.3, 0.5$ - but no relevant trend has been found. In Figure 4.14 we plot

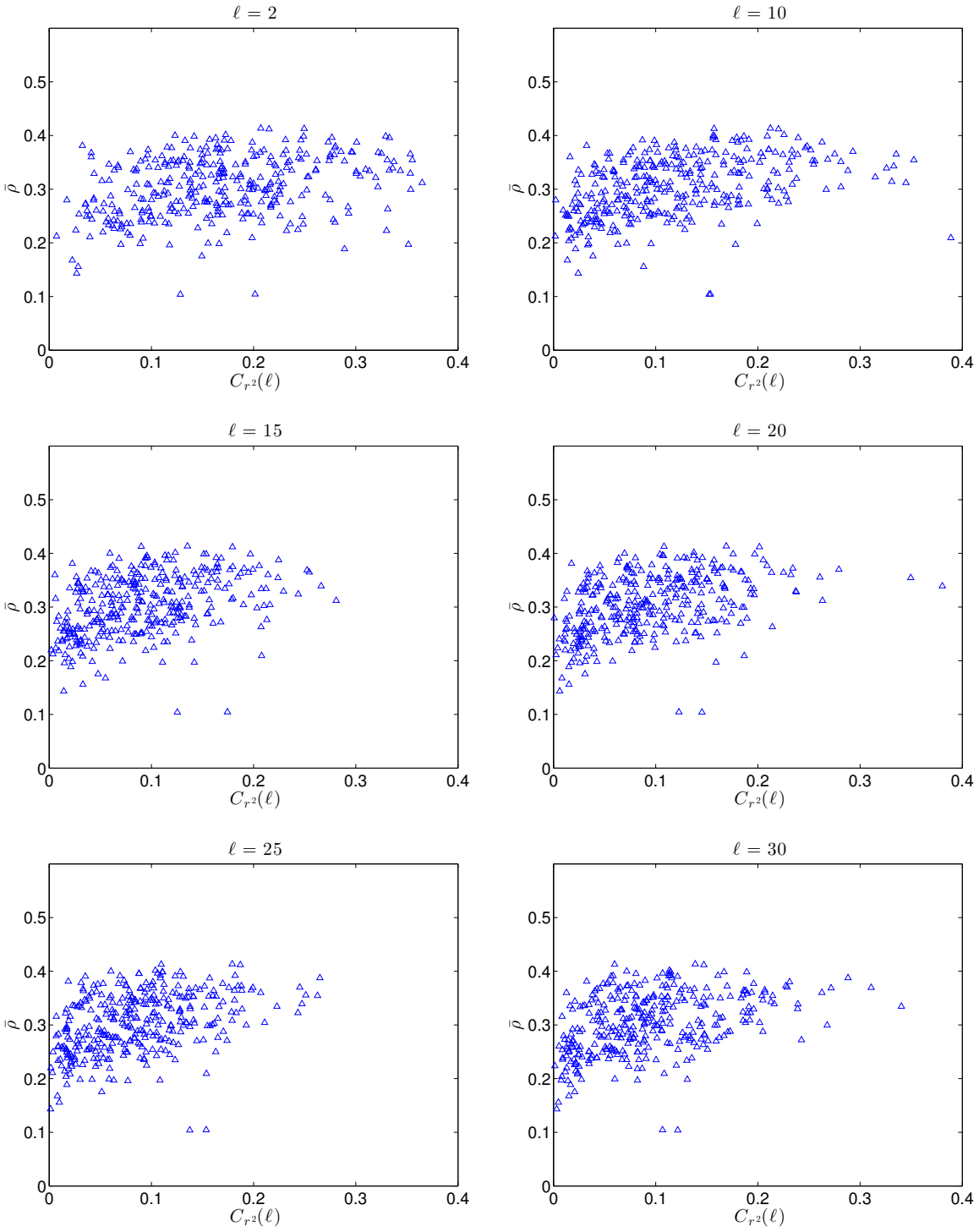

FIGURE 4.12: *Scatter plots of $\bar{\rho}_i$ vs $C_{r^2}(\ell)$ at several fixed ℓ's for all stocks for all stocks $i = 1, \ldots, N$.*

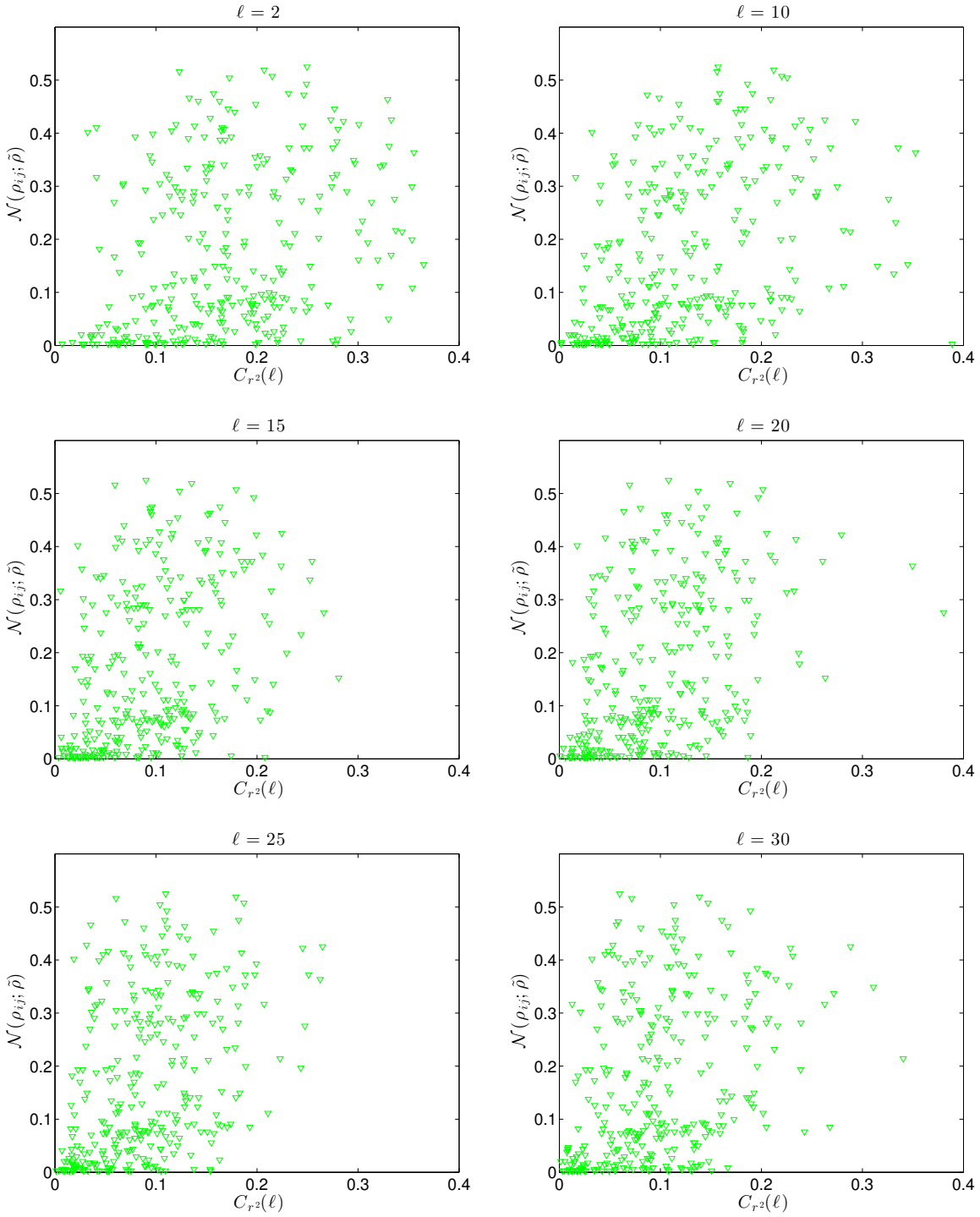

FIGURE 4.13: *Scatter plots of* $\mathcal{N}_i(\tilde{\rho} = 0.4)$ *vs* $C_{r^2}(\ell)$ *at several fixed* ℓ*'s for all stocks for all stocks* $i = 1, \ldots, N$.

	$\mathbf{C}_\alpha$		
ℓ	$\alpha = 1$	$\alpha = 2$	$\alpha = 3$
2	0.2695	0.2626	0.3473
3	0.4157	0.3416	0.5153
4	0.4503	0.4246	0.5057
5	0.4393	0.4462	0.4580
6	0.4370	0.3883	0.5207
7	0.4359	0.4394	0.5375
8	0.4804	0.4556	0.5410
9	0.4443	0.4499	0.4658
10	0.4278	0.4063	0.5385
11	0.4194	0.3776	0.4977
12	0.4657	0.4210	0.4949
13	0.4857	0.4565	0.5261
14	0.4140	0.4112	0.4387
15	0.3878	0.3946	0.4795
16	0.4607	0.4348	0.4743
17	0.4742	0.4430	0.5261
18	0.4895	0.4696	0.4886
19	0.4086	0.3784	0.5029
20	0.4157	0.4236	0.4665
21	0.4558	0.4435	0.4235
22	0.4624	0.4278	0.5224
23	0.4647	0.4455	0.5007
24	0.4621	0.4266	0.4995
25	0.4136	0.4217	0.4884
26	0.4278	0.4144	0.4778
27	0.4090	0.3892	0.5134
28	0.4852	0.4443	0.4979
29	0.4598	0.4500	0.4559
30	0.4036	0.3519	0.4976

TABLE 4.4: Values of $\mathbf{C}_\alpha$ for $\alpha = 1, 2, 3$ corresponding to $\ell \in [2, 30]$.

$\bar{\rho}_i^{PMFG}$ *vs* $C_{r^2}(\ell)$ for $\ell = 2, 10, 15, 20, 25, 30$. In this case the saturating behaviour observed for $\bar{\rho}$ is again present, particularly for $\ell > 10$, although the bending of the cloud of dots is less pronounced than in Figure 4.12.

We report in Table 4.4 the values of the correlation coefficients $\mathbf{C}_\alpha$ for $\alpha = 1, 2, 3$ and $\ell \in [2, 30]$. We also plot $\mathbf{C}_\alpha$ for $\alpha = 1, 2, 3$ against ℓ in figure 4.15. In all three cases $\mathbf{C}_\alpha$ attains its minimum value for $\tau = 2$, a feature which can be grasped intuitively and interpreted as follows: since small ℓ's correspond, on average, to the largest values of $C_{r^2}(\ell)$, the trend is necessarily flatter than that observed for larger ℓ's, as the larger $C_{r^2}(\ell)$ are plotted against the same array of $\mathcal{V}(\alpha)$ (which

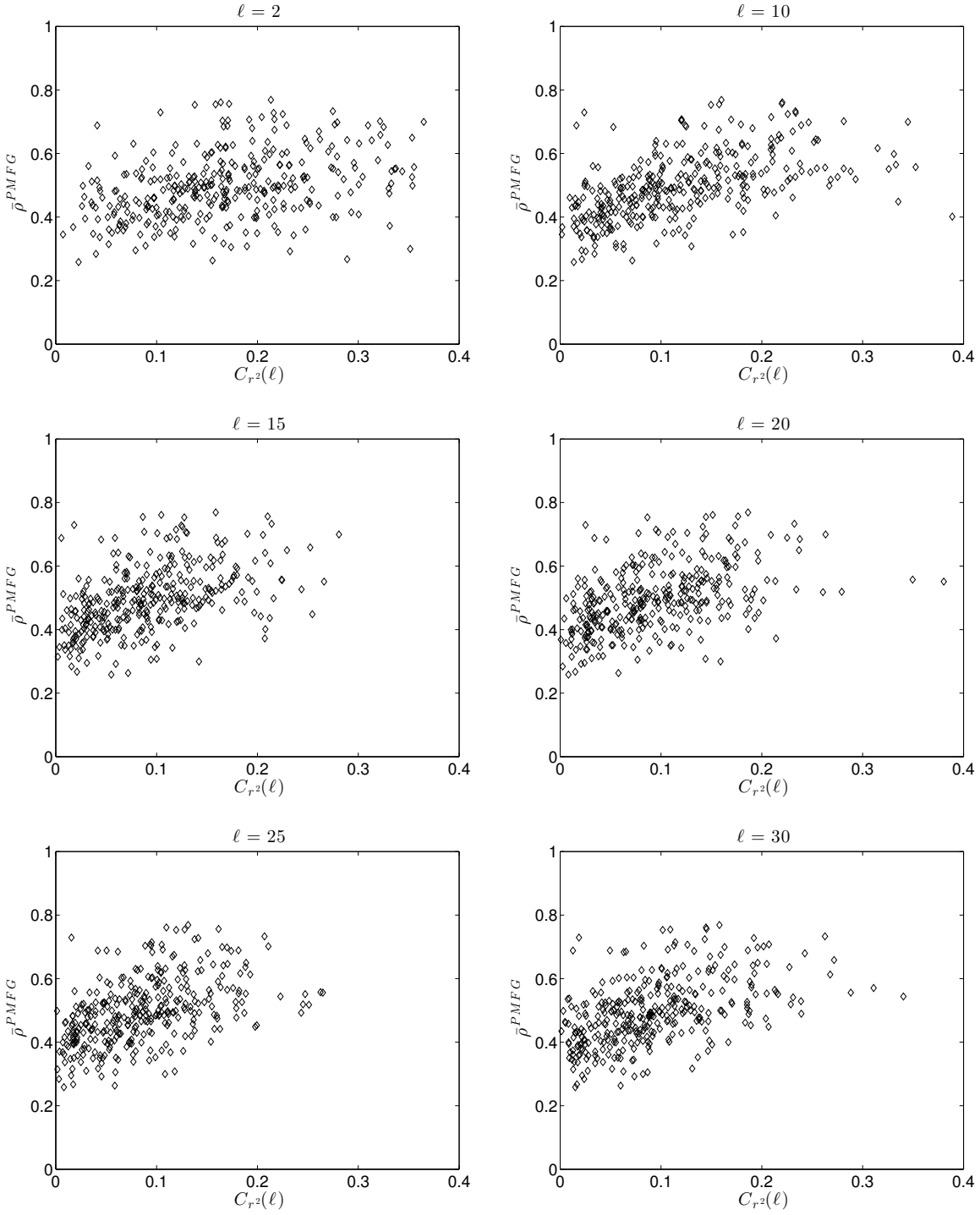

FIGURE 4.14: *Scatter plots of $\bar{\rho}_i^{PMFG}$ vs $C_{r^2}(\ell)$ at several fixed ℓ's for all stocks for all stocks $i = 1, \ldots, N$.*

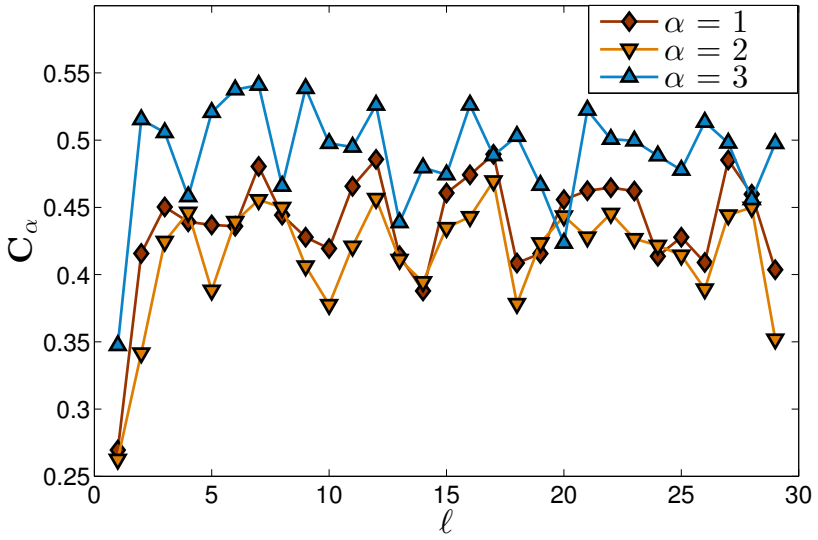

FIGURE 4.15: *Correlation* $\mathbf{C}_\alpha$ *between* $C_{r^2}(\ell)$ *and* $\mathcal{V}_\alpha$ *for* $\alpha = 1, 2, 3$. $\mathbf{C}_3$*, corresponding to the correlations filtered through the* $PMFG$ *is systematically larger than* $\mathbf{C}_2$ *and* $\mathbf{C}_1$*, suggesting that the correlations retained through the filtering procedure are those which show higher dependence with the volatility autocorrelation.*

doesn't change as we change ℓ).

We can see how, for $\ell > 4$, $\mathcal{V}_\alpha$ is positively correlated with $C_{r^2}(\ell)$ for all $\alpha = 1, 2, 3$. As was expected from Figure 4.13 $\mathbf{C}_\alpha$ is the lowest for $\alpha = 2$, but still fluctuates around 0.4, revealing that the two quantities cannot be considered completely independent from each other.

The most interesting feature that Figure 4.15 exhibits though, is that correlations filtered through the $PMFG$ are systematically more correlated to the volatility memory than the other two measures. In particular the average correlation $\bar{\rho}^{PMFG}$ is more correlated with $C_{r^2}(\ell)$ than the bare $\bar{\rho}$ is. This is intriguing as far as the relevance of the correlations filtered by the graph theoretic algorithm are concerned. Figure 4.15 would suggest that these filtered correlations also bear information about the volatility autocorrelations of the stocks involved.

This results therefore supports the relevance of the PMFG filtering procedure in identifying and retaining relevant components of the market and confirms the hypothesis of interplay between temporal and cross-correlation properties of stock returns.

4.5 Summary

In this chapter we have unveiled the interdependence between cross-correlation hierarchical properties and multifractality of stock returns. We have shown that, within a certain range of hierarchical orders, the degree multifractality is positively (although very likely non-linearly) dependent on the hierarchical order in the nested cross-correlation structure. This result has been found and investigated in great detail on NYSE data and further confirmed on other major international equity markets, LSE, TSE and HKSE. This analysis brings to light a new financial data stylised fact, which reveals further coherence in financial markets underlying organisation.

We have further studied linear correlations between cross-correlation coefficients and square returns autocorrelation function for fixed lags ℓ, confirming, through a different method, results recently established in [122]. We have further shown that correlation coefficients of stocks retained by the $PMFG$ are the most correlated with square returns autocorrelation functions. This results further highlights the relevance of the graph filtering technique in the identification of the most considerable part of the market.

Chapter 5

Multivariate Models for Stock Returns Dependency Structure

"If you give people literacy, bad ideas can be attacked and experiments tried, and lessons will accumulate."

- S.Pinker -

In this chapter we review multivariate models [23] commonly used to describe stock returns dependency structures and discuss multivariate stochastic volatility models [123] which incorporate, together with the cross-dependency structure, the temporal dependence observed in financial time series. We compare different models and provide an all-inclusive overview of the existing literature. By making use of existing formalism of Common Shock models [124], we further discuss a novel possible way of modelling a multivariate set of stocks incorporating the market hierarchical structure into a multivariate model with stochastic volatility.

5.1 Introduction

Modelling dependence in financial markets is one of the most relevant issues in quantitative finance and Econophysics. In every situation where many assets are

considered, the knowledge of the dependency structure governing the assets is of great importance for risk aversion, diversification and asset allocation [23]. The paramount example of the significance of the dependency structure in finance is portfolio optimisation, where one is interested in minimising the variance of the total portfolio of asset: the original solution found by Markowitz [125, 126], is that the optimal weights for the assets in the portfolio critically depend on the correlation matrix. The correct knowledge and estimation of the pair cross-correlations in financial markets is thus very important for an optimal investment. Cross-correlations between assets are important both within the same asset class (as is the case when dealing with a portfolio of equity for example) and between different asset classes: one may wish to invest in both stocks and bonds for example and it is therefore crucial to assess whether the two classes are strongly correlated or not in order to minimise the risk of loss by offsetting price falls in one class with increases in the other. The problem of describing the common movements in prices boils down to quantifying in probabilistic terms the distribution of joint movements. This problem is formalised by means of multivariate models, where a multi-dimensional probability distribution describes the joint behaviour of the entire portfolio of assets [127] (see Appendix F for a summary of the main concepts in multivariate statistics). Multivariate models are also widely employed in Risk Management to describe the joint distributions of different risk factors: suppose one wants to asses whether losses of the bank A are likely to be correlated with losses of another bank B: it is important to have a model describing what is the probability of observing these losses simultaneously and thus one needs to know how different risks are correlated between each other to avoid their occurrence all at once [128–131].

In this chapter we review the most relevant classes of multivariate models used in quantitative finance and risk management. These models can be divided into two main categories:

- **Static models**, i.e. models where the resulting multivariate time series of asset returns do not show any form of persistence.

- **Dynamic models**, incorporating a non-trivial temporal dependence together with the cross-correlation structure.

Among static models we will concentrate on the class of elliptical models [132] because of their relevance in finance: to this class of models belongs also the multivariate Gaussian model, which has enjoyed remarkable import in the modelling of correlated credit defaults [133], although according to most of the peers its adequacy has been unduly overrated [134]. In static models, although the unconditional multivariate distribution may have the desired properties observed in the real market, any time-dependence in returns is completely absent and therefore the resulting time series cannot reproduce the effect of volatility clustering observed in real data.

Static multivariate models can however be enhanced by adding a temporal structure to the time series and this is done through dynamic models: in this category we will discuss in Section 5.3 Multivariate GARCH (MGARCH) models [135] and Multivariate Stochastic Volatility (MSV) models [136]. MGARCH models generalise the univariate GARCH(p, q) of Bollerslev [86] to many assets and may reproduce also time-varying correlation depending on the model specification. MSV models instead are generalisation of static elliptical models, in that they attach all relevant stylised facts to the volatility, which is taken to be time correlated. Differently from static elliptical models though, MSV admit more than one volatility factor. In MGARCH models, volatility clustering and fat-tails are a consequence of a feedback mechanism, whereas in MSV models these stem from the choice of the statistical properties of the volatility. In this sense one may say that MGARCH models are more realistic as stylised facts emerge from a structural mechanism rather than from a particular parametric choice [137]. Nevertheless, although able to account for fat tails, the GARCH mechanism fails to reproduce long-range correlations in the volatility and one needs more than one exponential time scale for the power-law decay in the volatility autocorrelation function to appear [89, 90]. MSV instead can reproduce, together with fat tails, the desired long-range dependence by simply considering a power-law time-correlated volatility as in the MRW discussed in Chapter 3.

MSV models represent the best framework to reproduce a multivariate set of N cross-correlated as well as time-correlated time series. In spite of this, the specification of N different volatility components may be most of the time redundant, let alone the difficulty of handling the resulting large number of parameters. For this reason one often imposes a further factor structure [138, 139] to the MSV model, in order to reduce the dimension of the problem. The resulting picture is a MSV model with a reduced number of volatility factors, the simplest specification being that with one volatility factor only. This particular case is the closest one to the static elliptical models, the only difference being the temporal correlation in the volatility factor. In this class of models, we discuss in Section 5.4.1 the one factor Multivariate Multifractal Random Walk (MMRW) [24]. This model can reproduce, alongside fat tails and volatility clustering, multifractality of asset returns.
We finally investigate in Section 5.5 a novel multivariate model, based on Common Shock formalism [124]. This approach is endowed with a factor structure and it is shown to have a multi-time scale structure in the volatility autocorrelation. It requires nevertheless the specification of a further set of unobservable processes, which makes it very difficult to handle.

In the following we are going to review the models mentioned in this section. We start from static elliptical models, which lack any specification of the volatility dynamics and then discuss dynamic models.

5.2 Multivariate Elliptical Models

In this section we review static elliptical models. We first give general properties and definitions and then discuss in more details the three most notable examples: the multivariate Gaussian, the multivariate Student-t and the multivariate log-normal model.

Elliptical distributions are extremely popular to characterise the joint statistical

properties of many asset classes. They can be best understood as a more general case of another class of distributions, that of spherical distributions [23, 140].

Definition 1. [132] A random vector $\mathbf{X} = (X_1, \ldots, X_N)$ is spherically distributed if its distribution is invariant under any orthogonal map $\mathbf{\Gamma}$ (that is, any map which satisfies $\mathbf{\Gamma}\mathbf{\Gamma}^T = I_N$), i.e.

$$\mathbf{X} \stackrel{d}{=} \mathbf{\Gamma}\,\mathbf{X}, \tag{5.1}$$

where the notation $\stackrel{d}{=}$ stands for equality in distribution.

The fact that $\mathbf{X}$ has a spherical distribution is equivalent to the existence of a function $\psi : \mathbb{R} \to \mathbb{R}$ that, for any vector $a \in \mathbb{R}^N$, has the following property [132]

$$\psi(\mathbf{a}^T\mathbf{a}) = \psi(a_1^2 + \cdots + a_N^2) = \phi_{\mathbf{X}}(\mathbf{a}), \tag{5.2}$$

where $\phi_{\mathbf{X}}(\mathbf{a})$ is the characteristic function of $\mathbf{X}$ defined as

$$\phi_{\mathbf{X}}(\mathbf{a}) = \mathbb{E}(e^{i\mathbf{a}^T\mathbf{X}}). \tag{5.3}$$

The proof of the equivalence of these two definitions of spherical distributions can be found at page 89 of [132].
An alternative, very useful, representation of a spherical random vector is given by the following decomposition

$$\mathbf{X} \stackrel{d}{=} \sigma \cdot \mathbf{U}, \tag{5.4}$$

where $\mathbf{U}$ is a random vector uniformly distributed on the unit hypersphere $S_{N-1} \equiv \{\mathbf{x} \in \mathbb{R}^N : \ \mathbf{x}\mathbf{x}^T = 1\}$ in $\mathbb{R}^N$ and σ is a positive random variable independent of $\mathbf{U}$ [141]. Loosely speaking, a spherical random vector can be decomposed in a radial random component and a uniformly distributed multivariate random vector.

Definition 2. [132] A random vector $\mathbf{X}$ is elliptically distributed if

$$\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu} + A\mathbf{Y}, \tag{5.5}$$

where $\mathbf{Y}$ is spherically distributed, $\boldsymbol{\mu} \in \mathbb{R}^N$ is a vector of constants and $A \in \mathbb{R}^{N \times N}$ is a matrix such that $\boldsymbol{\Sigma} = AA^T$.

An elliptical random vector is thus obtained through an affine transformation of a spherical one. The vector $\boldsymbol{\mu}$ is usually referred to as the *location parameter*, whereas the covariance matrix $\boldsymbol{\Sigma}$ is called the *dispersion*. Another, more intuitive, way of characterising an elliptical random vector is through the iso-probability contours identified by the surfaces of the ellipsoids

$$\mathcal{E}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}^{C(L)}(\mathbf{x}) \equiv \{\mathbf{x}\colon \ (\mathbf{x}-\boldsymbol{\mu})\,\boldsymbol{\Sigma}^{-1}\,(\mathbf{x}-\boldsymbol{\mu})^T = C^2(L)\} \tag{5.6}$$

for $L \in (0, \infty)$. In two dimensions, i.e. for a bivariate distributions, these surfaces corresponds to ellipses in the (x_1, x_2) plane.
Starting from the representation (5.4), Embrechts *et al.* [127] show that elliptical random variables can be generated by multiplying a standardised Gaussian random vector $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \boldsymbol{\Sigma})$ with a random variable σ with support in $(0, +\infty)$ independent from $\boldsymbol{\epsilon}$:

$$\mathbf{X} = \boldsymbol{\mu} + \sigma\,\boldsymbol{\epsilon}. \tag{5.7}$$

The representation (5.7) immediately reminds us of the standard stochastic volatility representation discussed in Chapter 3, where a Gaussian residual is multiplied by a volatility process. The location parameter $\boldsymbol{\mu}$ is usually set to zero for financial returns, which can be assumed to be mean zero processes [27]. The distribution of $\mathbf{X}$ hence crucially depends on that of the volatility factor σ. The key point to remark is that whereas $\boldsymbol{\epsilon}$ is a multivariate vector, σ is the same for all stocks and thus can be interpreted as a common factor affecting all stocks. Chicheportiche and Bouchaud [142, 143] show that elliptical distributions fail to reproduce the multivariate dependency structure of stock returns especially when low-correlated stocks are considered. If one ranks the correlation coefficients ρ_{ij} for $i, j = 1, \ldots, N$, stocks exhibiting stronger discrepancies with elliptical models are those with $\rho < 0.4$ [142]; conversely, stocks exhibiting large correlation show better agreement with an elliptical dependency structure. These observations suggest that one single volatility factor is not enough to account for the complexity

of the market.

We now give details on the elliptical models most relevant to represent stock returns. These models are distinguished between each other via different choices of the volatility factor distribution $P(\sigma)$.

5.2.1 Multivariate Normal distribution

Multivariate Guassian random variables can be obtained by considering the volatility factor σ to be distributed as

$$P(\sigma) = \delta(\sigma - \bar{\sigma}), \tag{5.8}$$

where $\bar{\sigma}$ is the variance of $\mathbf{X}$. The resulting pdf of $\mathbf{X}$ is the multivariate normal distribution, with density [23]

$$f(\mathbf{x}) = \frac{1}{(2\pi)^N |\mathbf{\Sigma}|^{1/2}} \exp\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})' \mathbf{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\}. \tag{5.9}$$

In the special case of $\boldsymbol{\mu} = \mathbf{0}$ the density (5.9) is uniquely parametrised by the covariance matrix $\mathbf{\Sigma}$.

The bivariate Gaussian distribution is renown to provide a poor description of the dependency structure of stock returns [144], except for very low frequency (typically lower than the monthly returns). A possible way to assess multivariate normality is through Mardia's test [145, 146]. It is a test based on multivariate moments, namely skewness and kurtosis. One defines the following estimates for multivariate skewness and kurtosis

$$s_N = \frac{1}{n} \sum_{j=1}^{n} \sum_{i=1}^{n} D_{ij}^3, \qquad k_N = \frac{1}{n} \sum_{i=1}^{n} D_i^4, \tag{5.10}$$

where n is the length of the time series,

$$D_i^2 = (\mathbf{X}_i - \hat{\boldsymbol{\mu}}) \hat{\mathbf{\Sigma}}^{-1} (\mathbf{X}_i - \hat{\boldsymbol{\mu}}) \tag{5.11}$$

$$D_{ij} = (\mathbf{X}_i - \hat{\boldsymbol{\mu}}) \hat{\mathbf{\Sigma}}^{-1} (\mathbf{X}_j - \hat{\boldsymbol{\mu}}) \tag{5.12}$$

and the the hatted symbols stand for the estimated mean and covariance matrix, i.e.

$$\hat{\boldsymbol{\mu}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i, \qquad \hat{\boldsymbol{\Sigma}} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{X}_i - \hat{\boldsymbol{\mu}})(\mathbf{X}_i - \hat{\boldsymbol{\mu}})^T, \tag{5.13}$$

where all operations on vectors and matrices are to be performed componentwise. Under the null hypothesis of multivariate normality one expects in the limit $n \to \infty$ to have the two variables distributed as

$$\frac{1}{6}ns_N \sim \chi^2_{N(N+1)(N+2)/6}, \qquad \frac{k_N - N(N+2)}{\sqrt{8N(N+2)/n}} \sim \mathcal{N}(0,1). \tag{5.14}$$

The estimates (5.10) are then compared with the distributions (5.14) and the null hypothesis of multivariate normality is accepted if the p-value is above a certain reference value. We have performed the bivariate Mardia's test on a set of stocks' pairs arbitrary chosen from the 350 most capitalised quoted in NYSE in the period 1997-2012, fixing the acceptance threshold at 0.025. We show in Table 5.1 the results of the test: we report, for different time scales the p-values for both Skewness and Kurtosis. One can see that the hypothesis of bivariate normalty is rejected (p-value smaller than 0.025) for all pairs at daily, weekly and monthly (except one pair) frequencies, whereas cannot be rejected for many pairs when quarterly returns are considered. These results are in line with what already observed on different data sets [132]. The inadequacy of the bivariate Gaussian to describe the joint distribution of stock returns, particularly at daily frequency, is mainly due to its tails being too thin to account for joint large events, which are instead very often observed in financial time series [147].

5.2.2 Multivariate Student-t distribution

If the distribution of the square volatility σ^2 is chosen as inverse Gamma, which is equivalent to saying that $\sigma \sim \sqrt{\mathrm{Ig}(\nu/2, \nu/2)}$, i.e.

$$P(\sigma^2 = u) = \frac{1}{\Gamma(\frac{\nu}{2})}u^{-\frac{\nu}{2}-1}e^{-\frac{1}{u}}, \tag{5.15}$$

TABLE 5.1: P-values for Mardia's test of bivariate normality based on Skewness and Kurtosis performed on 10 pairs of stocks. We report the p-values for different returns frequencies. n denotes the number of observations for each time scale. The bold numbers are the p-values for which the hypothesis of joint normality cannot be rejected at a confidence level of 0.025.

Pairs	p-value		p-value	
	Skewness	Kurtosis	Skewness	Kurtosis
	Daily returns, n=4026		Weekly returns, n=805	
FHN,AGN	0.00	0.00	0.00	0.00
FHN,HON	0.00	0.00	0.00	0.00
BWS,BAX	0.00	0.00	0.00	0.00
BDX,BWS	0.00	0.00	0.00	0.00
ABX,AFL	0.00	0.00	0.00	0.00
AFL,AEP	0.00	0.00	0.00	0.00
ALTR,HES	0.00	0.00	0.00	0.00
ALTR,AIG	0.00	0.00	0.00	0.00
ADM,AXP	0.00	0.00	0.00	0.00
ADM,AM	0.00	0.00	0.00	0.00
	Monthly returns, n=192		Quarterly returns, n=48	
FHN,AGN	6.66×10^{-16}	0.00	**0.1451**	0.0065
FHN,HON	0.00	0.00	**0.1672**	0.0083
BWS,BAX	8.07×10^{-11}	0.00	**0.2994**	**0.1354**
BDX,BWS	1.98×10^{-9}	0.00	0.0215	**0.0751**
ABX,AFL	0.00	0.00	0.0263	3.16×10^{-7}
AFL,AEP	0.00	0.00	1.53×10^{-11}	3.33×10^{-16}
ALTR,HES	9.43×10^{-13}	0.00	**0.1516**	**0.1580**
ALTR,AIG	6.88×10^{-8}	0.00	**0.2078**	**0.2025**
ADM,AXP	1.35×10^{-13}	0.00	**0.2871**	**0.4277**
ADM,AM	**0.1228**	0.00	0.0154	0.0185

then $\mathbf{X}$ has a multivariate Student-t distribution with ν degrees of freedom, $\mathbf{X} \sim \text{St}(\nu, \boldsymbol{\mu}, \boldsymbol{\Sigma})$. The probability density for a multivariate Student-t vector can be shown to be [141, 148, 149]

$$f(\mathbf{x}) = \frac{\Gamma(\frac{1}{2}(\nu+N))}{\Gamma(\frac{\nu}{2})(\pi\nu)^d|\boldsymbol{\Sigma}|^{1/2}} \left(1 + \frac{(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}{\nu}\right)^{-(\nu+N)/2}. \tag{5.16}$$

The Student-t distribution enjoys a set of properties which make it appealing to model stock returns:

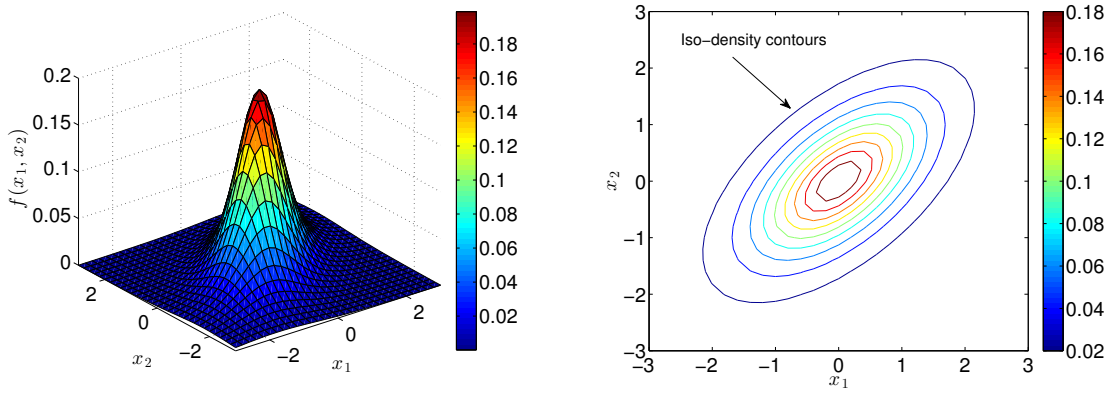

FIGURE 5.1: Probability density function (left) and iso-density contours (right) for a bivariate Student-t distribution with $\nu = 3.5$ and $\Sigma_{12} = 0.4$.

- It has thick tails that can account for large common movements in the variables $(X_1, \ldots, X_N)$, a feature which was missed by the multivariate Gaussian. More specifically it has a much pronounced leptokurtic shape, with contours of equal probability density much steeper around the centre of the distribution than in the peripheral areas. We plot in Figure 5.1 the empirical pdf and contour plot obtained for a bivariate Student-t random vector with $\nu = 3.5$ and $\Sigma_{12} = 0.4$: both panels show the leptokurtic behaviour.

- The covariance matrix is given by $\text{Cov}(\mathbf{X}) = \frac{\nu}{\nu-2}\mathbf{\Sigma}$ [23].

- The marginal distributions of each of the X_i for $i = 1, \ldots, N$ are Student-t distributions with ν degrees of freedom and variances given by the diagonal elements of the matrix $\text{Cov}(\mathbf{X})$.

- Any affine transformation $\mathbf{Z} = \mathbf{a} + \mathbf{B}\mathbf{X}$ is still Student-t distributed, according to

$$\mathbf{Z} \sim \text{St}(\nu, \mathbf{a} + \mathbf{B}\boldsymbol{\mu}, \mathbf{B}\mathbf{\Sigma}\mathbf{B}^T).$$

- In the limit $\nu \to \infty$ the multivariate Student-t distribution boils down to a multivariate Gaussian, $\text{St}(\nu, \boldsymbol{\mu}, \mathbf{\Sigma}) \overset{\nu\to\infty}{\to} \mathcal{N}(\boldsymbol{\mu}, \mathbf{\Sigma})$, with the tails of the joint density growing thicker as the degrees of freedom are decreased.

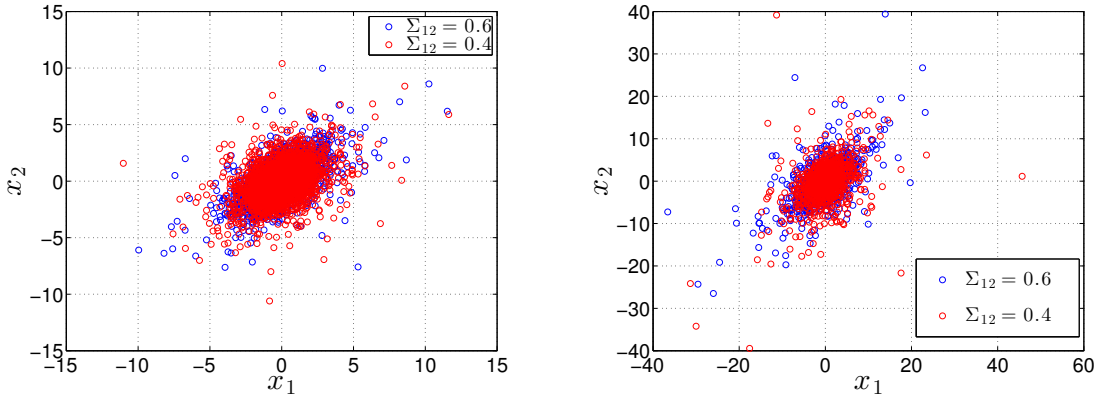

FIGURE 5.2: (Left) Scatter plot of two bivariate Student-t vectors $\mathbf{X} = (X_1, X_2)$ with $\Sigma_{12} = 0.4$ (red circles) and $\Sigma_{12} = 0.6$ (blue circles) and $\nu = 4$. (Right) Scatter plot of two bivariate log-normal vectors $\mathbf{X} = (X_1, X_2)$ with $\Sigma_{12} = 0.4$ (red circles) and $\Sigma_{12} = 0.6$ (blue circles) and $\omega_0 = 1$.

We show in the left panel of Figure 5.2 a scatter plot of a bivariate Student-t random vector of length $n = 4000$, for two different values of the covariance coefficient, $\Sigma_{12} = 0.4$ (red circles) and $\Sigma_{12} = 0.6$ (blue circles) respectively.

5.2.3 Multivariate Log-Normal distribution

Another relevant model which has received much attention in volatility modelling, is the log-normal, particularly on studies on multi fractals [74, 88]. The multivariate distribution is obtained when we choose the volatility factor σ to be distributed as $\sigma \sim e^{\omega}$, with $\omega \sim \mathcal{N}(0, \omega_0)$. The resulting joint density is not available in closed form but can be shown to be (see Appendix G)

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{N/2}|\mathbf{\Sigma}|^{1/2}} \frac{1}{\sqrt{2\pi}\omega_0} \int_0^{\infty} \sigma^{-(\frac{\log\sigma}{2\omega_0^2}+\frac{N}{2}+1)} \exp\left(\frac{(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}{2\sigma}\right) d\sigma. \tag{5.17}$$

Note that this model must not be confused with the multivariate log-normal model, obtained by considering $\mathbf{X} \sim e^{\mathbf{Y}}$ with $\mathbf{Y} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{\Sigma})$[1]. In such case in fact, the vector $\mathbf{X}$ can assume only strictly positive values, which rule it out as a

[1]In this case the resulting multivariate distribution pdf reads [23]

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{N/2}\mathbf{\Sigma}^{1/2}\prod_{i=1}^{N} x_i} \exp\left(-\frac{1}{2}(\ln \boldsymbol{x} - \boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\ln \boldsymbol{x} - \boldsymbol{\mu})\right).$$

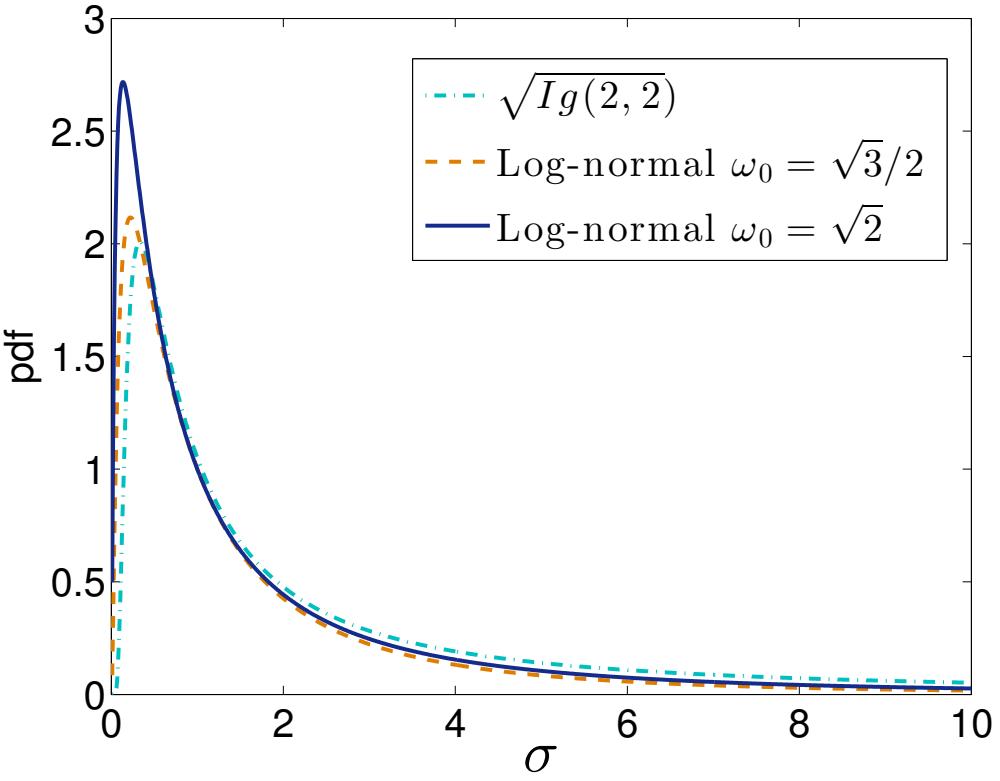

FIGURE 5.3: Comparison between log-normal pdf with two different values of ω_0 and $\sqrt{\text{Ig}(2,2)}$ pdf.

realistic candidate to model stock returns.

We show a scatter plot of a bivariate vector of length $n = 4000$ with log-normal volatility σ in the right plot of Figure 5.2 again for $\Sigma_{12} = 0.4$ (red circles) and $\Sigma_{12} = 0.6$ (blue circles) and $\omega_0 = 1$.

The multivariate log-normal and Student-t distributions provide a very similar description of stock returns dependency structure [143]. This is because the log-normal distribution shows remarkable similarities with the square root of the inverse gamma. In Figure 5.3 we plot the pdf of an inverse-gamma distributed random variable under square root, compared with the pdf of two log-normal random variables with $\omega_0 = \sqrt{3}/2$ and $\omega_0 = \sqrt{2}$ respectively. Especially in the tail region, it is by all practical purposes nearly impossible to distinguish one pdf from the other. This results in the two models (the Student-t and the log-normal) having very similar statistical properties.

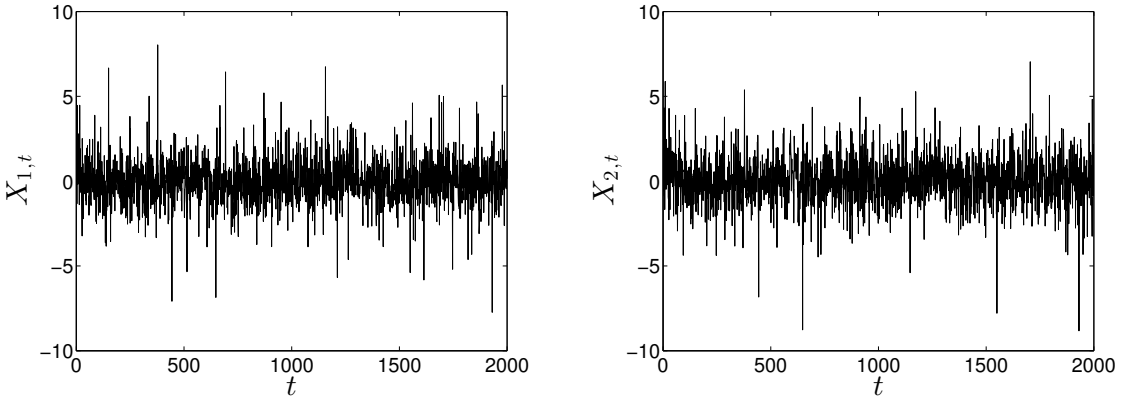

FIGURE 5.4: Realisations of a bivariate Student-t random vector $\mathbf{X}_t = (X_{1,t}, X_{2,t})$ with $\nu = 4$ and $\Sigma_{12} = 0.6$, lacking any relevant temporal dependence.

5.3 Multivariate models with stochastic volatility

In static elliptical models we have only considered the cross-dependency structure between different stocks to be the relevant feature of the different multivariate models. The volatility factor σ has been taken as a stochastic process whose time components $\{\sigma_t,\ t = 1, \ldots, n\}$ were i.i.d. random variables. This specification is clearly not optimal to realistically describe the dynamics of financial time series, that, as we have extensively discussed in Chapter 2, show remarkable persistence and more precisely heteroskedastic volatility. A realisation of a bivariate Student-t random vector described in 5.2.2 for instance, yields returns that clearly do not show any temporal dependence (see Figure 5.4), although they still may capture stock returns cross-dependency structure. The static multivariate elliptical models seen so far cannot account for any serial dependence in the common movements of asset returns, which are of the uttermost importance in risk management and optimal asset allocation problems [132]. To overcome this limitation one resorts to multivariate models with stochastic volatility. As for what seen with univariate models, the volatility is not only let depend on time, but is also given a non-trivial temporal dependence, which can vary according to the model specification. The

first and largest class of multivariate models with time correlations is that of autoregressive conditional variance models, more generally referred to as multivariate GARCH (MGARCH) models.

5.3.1 Multivariate GARCH models

A multivariate GARCH process is a strictly stationary process $\mathbf{X}_t$ defined by [135]

$$\boldsymbol{X}_t = \boldsymbol{\Sigma}_t^{1/2} \boldsymbol{Z}_t \tag{5.18}$$

where $\boldsymbol{Z}_t \sim \mathcal{N}(\mathbf{0}, \mathbb{I}_N)$ and the covariance matrix $\boldsymbol{\Sigma}_t$ is time-dependent. The simplest specification of $\boldsymbol{\Sigma}_t$ is the one which defines the Constant Conditional Correlation GARCH (CCC-GARCH), satisfying the equations [150]

$$\begin{aligned} \boldsymbol{\Sigma}_t &= \Delta_t C \Delta_t \\ \Delta_t &= \mathrm{diag}(\sigma_{1,t}, \ldots, \sigma_{N,t}) \\ \sigma_{i,t} &= \alpha_{i,0} + \sum_{j=1}^{p_i} \alpha_{i,j} X_{i,t-1}^2 + \sum_{k=1}^{q_i} \beta_{i,k} \sigma_{i,t-k}^2, \qquad i = 1, \ldots, N \end{aligned} \tag{5.19}$$

where C is a constant positive-definite correlation matrix, $\alpha_{i,0} > 0$, $\alpha_{i,j} \geq 0$ for $j = 1, \ldots, p_i$ and $\beta_{i,k} \geq 0$ for $k = 1, \ldots, q_i$. The model specified in Equations (5.19) has $N(N+5)/2$ parameters, if one assumes that $p_i = q_i = 1$, $\forall i = 1, \ldots, N$. The covariance matrix $\boldsymbol{\Sigma}_t$ is positive-definite, thanks to the fact that C is positive-definite[2]. As in the univariate case, the variances $\sigma_{i,t}$ for $i = 1, \ldots, N$ evolve trough the GARCH feedback mechanism but now returns of different stocks are also cross-correlated. Differently from the models discussed in Section 5.2, the multivariate structure is attached to the volatility factor, which is incorporated into the covariance matrix $\boldsymbol{\Sigma}_t$, while the Gaussian residuals are uncorrelated both in time and cross-sectionally. The univariate time series show the usual GARCH

[2]To prove a matrix M is positive-definiteness one needs to check that, for any strictly positive vector $\boldsymbol{v}$ the condition $\boldsymbol{v}^T M \boldsymbol{v}$ is enforced. In the CCC-GARCH case we have $\boldsymbol{v}^T \boldsymbol{\Sigma}_t \boldsymbol{v} = (\Delta_t \boldsymbol{v})^T C (\Delta_t \boldsymbol{v}) > 0$ as C is positive-definite and $\Delta_t \boldsymbol{v} \neq 0$ because of the non-negativity of the volatility processes.

properties, namely fat tails of the unconditional returns distribution and exponential decay of the volatility autocorrelation function for each $i = 1, \ldots, N$.

In the CCC-GARCH the correlation matrix is conditionally constant, a feature which is generally regarded as too restrictive, given the sweeping empirical evidence of time-varying correlation [1, 44, 120, 151–154]. A more advanced model allowing conditional correlations to evolve in time is the Dynamical Conditional Correlation GARCH (DCC-GARCH) [155, 156]. In the DCC-GARCH the univariate volatilities still evolve as in Equation (5.19), but the constant correlation matrix C is replaced by a time-dependent correlation matrix C_t that satisfies the following equation:

$$C_t = \mathcal{P}\left(\left(1 - \sum_{j=1}^{p} \alpha_j + \sum_{k=1}^{q} \beta_k\right) C + \sum_{j=1}^{p} \alpha_j \boldsymbol{Y}_{t-1} \boldsymbol{Y}_t^T + \sum_{k=1}^{q} \beta_k C_{t-1}\right) \tag{5.20}$$

where $\mathcal{P}(\cdot)$ is a matrix operator which acts on a matrix M as

$$\mathcal{P}(M) = \operatorname{diag}(\sqrt{M_{11}}, \ldots, \sqrt{M_{NN}})^{-1} M \operatorname{diag}(\sqrt{M_{11}}, \ldots, \sqrt{M_{NN}})^{-1}, \tag{5.21}$$

the coefficients α_j and β_k are non-negative and satisfy $\sum_{j=1}^{p} \alpha_j + \sum_{k=1}^{q} \beta_k < 1$ and $\boldsymbol{Y}_t = \Delta_t^{-1} \boldsymbol{X}_t$. If all α_i and β_i are zero the correlation matrix (5.20) reduces to C and the DCC model boils down to the CCC-GARCH. For $p = q = 1$ the DCC-GARCH has $(N + 1)(N + 4)/2$ parameters. Note that the covariance matrix of DCC-GARCH is also positive-definite. Let us just replace the operator $\mathcal{P}(\cdot)$ with the identity, since dividing by the positive square roots of the variances would not affect the positive-definiteness. If we define

$$C_t^I = \left(1 - \sum_{j=1}^{p} \alpha_j + \sum_{k=1}^{q} \beta_k\right) C + \sum_{j=1}^{p} \alpha_j \boldsymbol{Y}_{t-1} \boldsymbol{Y}_t^T + \sum_{k=1}^{q} \beta_k C_{t-1}, \tag{5.22}$$

then, if we assume $C_{t-q} \dots C_{t-1}$ to be positive-definite, for any positive vector $\boldsymbol{v}$ one has $\boldsymbol{v}^T C_t^I \boldsymbol{v} > 0$. The three terms to consider are indeed

$$(1) \qquad \left(1 - \sum_{j=1}^{p} \alpha_j + \sum_{k=1}^{q} \beta_k\right) \boldsymbol{v}^T C^I \boldsymbol{v}$$

$$(2) \qquad \sum_{j=1}^{p} \alpha_j \boldsymbol{Y}_{t-1} \boldsymbol{v}^T \boldsymbol{Y}_t^T \boldsymbol{v}$$

$$(3) \qquad \sum_{k=1}^{q} \beta_k \boldsymbol{v}^T C_{t-1} \boldsymbol{v}.$$

Since (1) is strictly positive and (2) and (3) are non-negative C_t^I is positive-definite and so is C_t. The general criticism [123, 135] toward the DCC specification is that the entire covariance matrix shares the same parameters α and β, which, especially when considering a large number of stocks, is altogether unrealistic. For this reason it has also been proposed to allow α and β to be N-dimensional vectors while preserving the positive-definiteness of both C_t and C_t^I [157]. This adds richness to the model but pushes up of $2N$ the number of parameters.

A further step in complexity, although forerunning the DCC-GARCH, in the family of MGARCH models corresponds to specify the conditional variance $\boldsymbol{\Sigma}_t$ explicitly. The most considerable models belonging to this class are the DVEC model [158] and the Baba, Engle, Kroner and Kraft (BEKK) model [159]. Whereas the first one is a constrained version of the more general but over-parametized VEC model [159], the second ensures the positive-definiteness of the covariance matrix without imposing further conditions as for the DVEC. Because of their very large number of parameters though ($(5N^2 + N)/2$ for BEKK and $3N(N + 1)/2$ for DVEC) these models can be used only for very low-dimensional sets of stocks and were progressively put aside, the DCC specification being much more flexible.

5.3.2 Multivariate stochastic volatility models

A different approach to model a multivariate set of stocks with temporal correlation is through multivariate stochastic volatility models. Differently from the MGARCH setting, the covariance matrix does not depend on past observations

but on unobserved volatility processes. In the pioneering study [160] on MSV models, returns r_t are modelled as

$$\boldsymbol{r}_t = H^{1/2}\boldsymbol{\xi}_t \tag{5.23}$$

$$H_t^{1/2} = \text{diag}(\exp(h_{1,t}), \ldots, \exp(h_{N,t})) \tag{5.24}$$

$$h_{i,t} = \phi_i h_{i,t-1} + \eta_{i,t}, \qquad i = 1, \ldots, N \tag{5.25}$$

with $\boldsymbol{\xi}_t \sim \mathcal{N}(0, \Sigma_\xi)$ and $\boldsymbol{\eta}_t \sim \mathcal{N}(0, \Sigma_\eta)$, having denoted by Σ_ξ and Σ_η the two covariance matrices. The model in Equations (5.23)-(5.25) has $N^2 + 2N$ parameters. The univariate log-volatilities in Equation (5.24) are AR(1) processes so that the volatility components of $H_t^{1/2}$ are conditionally log-normal. Note that this model accounts for cross-correlation both between returns residuals $\boldsymbol{\xi}_t$ and volatility innovations $\boldsymbol{\eta}_t$, hence providing a richer structure with respect to MGARCH. The CCC-GARCH setting can be recovered if one sets to zero all off-diagonal entries of Σ_η. The model proposed in [160] has also been extended to account for fat tails, specifically considering Student-t distributed residuals ξ_t, which can reproduce larger values for the excess kurtosis with respect to the original Gaussian model, in agreement with empirical observations. Another extension concerns the introduction of returns-volatility correlations which can account for both asymmetries and leverage effect [136]. These two concepts are very often confused although they refer to two completely different properties observed in financial time series: the asymmetry of volatility refers to the different effect that positive and negative returns have on volatility, whereas the leverage effect denotes the negative correlation observed between current returns and future volatility. A possible extension of the model of [160] which can account for leverage effect is found in [161] starting from an idea first proposed in [162], where the multivariate structure has the following form

$$\begin{pmatrix} \boldsymbol{\xi}_t \\ \boldsymbol{\eta}_t \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_\xi & L \\ L & \Sigma_\eta \end{pmatrix} \right]$$

$$L = \text{diag}(\lambda_1 \sigma_{\eta,11}, \ldots, \lambda_N \sigma_{\eta,NN}), \tag{5.26}$$

the $\sigma_{\eta,ii}$ for $i = 1, \ldots, N$ are the diagonal entries of Σ_η and the parameters λ_i for $i = 1, \ldots, N$ are expected to be negative. The model (5.26) has $3N + N^2$ parameters.

5.4 Factor models

To reduce the dimension of the space of parameters of a model, one often resorts to factor models [138, 139]. Factor models are justified intuitively by the common knowledge that the randomness of a set of stocks can be explained in terms of a smaller number of factors, shared by subsets of stocks. The most straightforward example is that of a pool of equity, which, although only to a first approximation, can be described by means of the market index including them. The most general d-factor model expresses the N-dimenisonal random vector $\mathbf{X}$ as

$$\mathbf{X} = W \cdot \mathbf{F} + \boldsymbol{\eta} \tag{5.27}$$

where $\mathbf{F}$ is a d-dimensional vector of factors with $d < N$, $\boldsymbol{\eta}$ is a d-dimensional vector of mean-zero idiosyncratic noises uncorrelated among them and uncorrelated with $\mathbf{F}$ and W is a $N \times d$ matrix of constant factor loadings. The dynamics of the component X_i is therefore explained through the common action of different factors, weighted by the coefficients w_{ij}, with $j = 1, \ldots, d$. The simplest situation is the so-called one-factor model, where only one factor is responsible for the dynamics of the entire pool of stocks:

$$\mathbf{X} = W \cdot F + \boldsymbol{\eta} \tag{5.28}$$

where F now is a random variable and $W = (w_1, \ldots, w_N)$ embodies the impact that the factor F has on different stocks. If we denote by σ_F and σ_{η_i} the variances of the common factor F and of η_i the correlation matrix of the one-factor model can be easily shown to have entries [27, 163]

$$C_{ij} = w_i w_j \sigma_F^2 + \sigma_{\eta_i}^2 \delta_{ij}, \tag{5.29}$$

where δ_{ij} denotes the Kronecker delta. The model in Equation (5.28) is clearly over-simplistic to describe the complexity of correlation in the market where typically sectors have a relevant influence on the cross-correlation between stocks [107, 164].

Factor models have been employed in the context of MSV models to reduce the number of independent variables and thus simplify the structure of the time series modelling problem. We mention that they are used to the same aim also for MGARCH models, but we won't discuss this topic here. There are mainly two types of dimension reduction through factor models that are particularly relevant for MSV models: additive factor models and multiplicative factor models. Additive factor models for MSV models were first introduced in [160] and are discussed further in [136]. They are built combining Equation (5.27) with the following dynamics of the factors $(F_1, \dots, F_d)$:

$$\begin{aligned} F_{i,t} &= \exp(h_{i,t}/2)\xi_{i,t} \\ h_{i,t} &= \phi_i h_{i,t-1} + \sigma_{\gamma_i}\gamma_{i,t}, \qquad i = 1, \dots, d \end{aligned} \tag{5.30}$$

where $\xi_{i,t} \sim \mathcal{N}(0,1)$ and $\gamma_{i,t} \sim \mathcal{N}(0,1)$ for $i = 1, \dots, d$. Yu and Meyer [165] show, in the simplified scenario of a bivariate one-factor model, that the specification (5.30) can account for both time varying volatility and correlation and that correlation increases along with the volatility of the factor. The major drawback of this model is however the following: since the idiosyncratic errors are not heteroskedastic, if the number of factor is smaller than the number of stocks, one can show that there will always be stocks whose prices show no heteroskedasticity, a feature which clashes with empirical stylised facts [136].

Multiplicative factor models instead, considered for example in [166], are the closest type of model to the static multivariate models discussed in Section 5.2. Their simplest formulation is a multiplicative one-factor model, where returns are decomposed into a common volatility factor and a vector of idiosyncratic noise in

the following way

$$\begin{aligned} \boldsymbol{r}_t &= \exp(h_t/2)\boldsymbol{\xi}_t, \qquad \boldsymbol{\xi}_t \sim \mathcal{N}(0, \Sigma_\xi) \\ h_t &= \mu + \phi(h_{t-1} - \mu) + \eta_t, \qquad \eta_t \sim \mathcal{N}(0, \sigma_\eta) \end{aligned} \tag{5.31}$$

where h_t and η_t are now scalar random variables and μ and ϕ are constants. Multi-factor multiplicative models can be then obtained replacing the volatility factor $\exp(h_t/2)$ by the vector

$$\boldsymbol{\sigma}_{W,t} = \exp\left(\frac{W \cdot \boldsymbol{h}_t}{2}\right) \tag{5.32}$$

where W is an $N \times d$ matrix of factor loadings and h_t is now a d-dimensional vector. The AR(1) dynamics for h_t can also be replaced by other more complex ones, particularly if one wants to account for long-ranged dependence in the volatility. For this reason we can reformulate in a more general way the models in Equations (5.31) and (5.32) through the following equations:

$$\boldsymbol{r}_t = \exp(\omega_t/2)\boldsymbol{\epsilon}_t \qquad \text{one-factor} \tag{5.33}$$

$$\boldsymbol{r}_t = \exp\left(\frac{W \cdot \boldsymbol{\omega}_t}{2}\right)\boldsymbol{\epsilon}_t \qquad \text{multi-factor} \tag{5.34}$$

where ω_t is a random variable with auto-covariance function $C_\ell = \mathrm{Cov}(\omega_t \omega_{t+\ell})$ such that

$$\int_0^\infty C_\ell \, d\ell = \infty, \tag{5.35}$$

$\boldsymbol{\omega}_t$ is a d-dimensional random vector of volatilities for which Equation (5.35) holds componentwise and $\boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \Sigma_\xi)$. The one-factor multiplicative model in Equation (5.33) is very appealing in its formulation because of its parsimony. It also separates the temporal dependence, carried by the common volatility factor, from the cross-sectional dependence, embodied in the multivariate vector of residuals $\boldsymbol{\epsilon}_t$. The multi-factor multiplicative model (5.34) further introduces d volatility factors which affect the N stocks according to the weighting matrix W. In the general class of models of Equation (5.34), a particularly relevant one in the context of this thesis is the multivariate Multifractal Random Walk (MMRW).

5.4.1 Multivariate Multifractal Random Walk

The MMRW, introduced in [167], extends the univariate MRW to describe the dynamics of many assets relevant for portfolio management [24]. It can be recovered from Equation (5.34) by considering $d = N$ volatility factors and a matrix $W = \mathbb{I}_N$ such that, at scale Δt, the process increments $\boldsymbol{r}_{t,\Delta t} = (r^{(1)}_{t,\Delta t}, \ldots, r^{(N)}_{t,\Delta t})$ are

$$\boldsymbol{r}_{t,\Delta t} = \exp(\boldsymbol{\omega}_{t,\Delta t}) \cdot \boldsymbol{\epsilon}_{t,\Delta t} \tag{5.36}$$

where the process $\boldsymbol{\epsilon}_{t,\Delta t}$ is a multivariate Gaussian vector with covariance

$$\mathrm{Cov}(\epsilon^{(i)}_{t,\Delta t}\epsilon^{(j)}_{t+\tau,\Delta t}) = \delta_{\tau 0}\, \boldsymbol{\Sigma}_{ij}\, \Delta t. \tag{5.37}$$

The Kronecker delta $\delta_{\tau 0}$ excludes any lagged cross-correlation and $\boldsymbol{\Sigma}_{ij}$ quantifies the cross-correlations between the different residuals $\epsilon^{(i)}_{t,\Delta t}$. The magnitude vector $\omega_{t,\Delta t}$ has covariance

$$\mathrm{Cov}(\omega^{(i)}_{t,\Delta t}\omega^{(j)}_{t+\tau,\Delta t}) = \begin{cases} \boldsymbol{\Lambda}_{ij} \ln \frac{T_{ij}}{\Delta t+\tau}, & t+\tau < T_{ij} \\ 0, & \text{elsewhere} \end{cases} \tag{5.38}$$

with $\boldsymbol{\Lambda}_{ij}$ being a matrix denoted *multifractal matrix* whose diagonal elements are the intermittency coefficients λ_i^2 of the N volatility factors ω_i for $i = 1 \ldots, N$. The limit $\Delta t \to 0$ of the cumulative sum process $X_{t,\Delta t} = \sum_{k=1}^{t/\Delta t} r_{k,\Delta t}$ is a multifractal process obeying the following scaling law

$$\mathbb{E}(|r^{(1)}_{t,\Delta t}|^{q_1} \ldots |r^{(N)}_{t,\Delta t}|^{q_N}) = K_{q_1 \ldots q_N} \Delta t^{\zeta_{q_1 \ldots q_N}} \tag{5.39}$$

with $K_{q_1 \ldots q_N}$ a constant depending on $q_1 \ldots q_N$. Bacry *et al.* discuss in [167] how, regardless of the form of $\boldsymbol{\Sigma}_{ij}$, the multifractal spectrum is

$$\zeta_{q_1 \ldots q_N} = \sum_{i=1}^{N} \zeta^i_{q_i} \; - \sum_{1 \le i \le j \le N} \boldsymbol{\Lambda}_{ij} q_i q_j. \tag{5.40}$$

5.4.2 One-factor MMRW

The one-factor version of the model discussed in 5.4.1 can be defined by considering only one volatility factor, common to all N stocks. In this case the volatility process $\sigma_{t,\Delta t}$ is the same as defined in Chapter 2 and the only difference with the univariate MRW is in the Gaussian residual factor being a multivariate vector $\boldsymbol{\epsilon}_{t,\Delta t}$. The return vector is

$$\boldsymbol{r}_{t,\Delta t} = \exp(\omega_{t,\Delta t})\boldsymbol{\epsilon}_{t,\Delta t}. \tag{5.41}$$

If one identifies the generic random vector $\boldsymbol{X}$ of Section 5.2 with $\boldsymbol{r}_{t,\Delta t}$ then the one-factor MMRW represents a concise generalisation of the multivariate model of Section 5.2.3 with a non-trivial temporal dependency structure. In this specification the entire set of stocks has only one volatility mode, whereas all idiosyncrasies are quantified in the vector of residuals $\boldsymbol{\epsilon}_{t,\Delta t}$. The model can be modified to account for thicker tails by considering $\boldsymbol{\epsilon}_{t,\Delta t}$ to be a Student-t multivariate vector as in Section 5.2.2.

5.5 Common Shock Multivariate Models

In this section we investigate a novel alternative way to model a multivariate set of stocks. We make use of the existing formalism of Common Poisson Shock (CPS) which, to the best of our knowledge at the draft of this thesis, have never been used before to model stock returns' dynamics. We modify the original setting of CPS introducing Hawkes processes as the underlying point processes, showing that this can produce long-range correlation in the volatility process. This approach is also shown to naturally envisage a hierarchy in the differente risks affecting the time series.

5.5.1 Common Poisson Shock formalism

The Common Poisson Shock Model (CPS), discussed extensively in [124], has been used both in reliability theory and in insurance and credit risk modelling [168]. In this paragraph we briefly review, on the lines of [124], the main definitions which we will use as a setup for the Common Hawkes Shocks model (CHS), discussed in the next section. The basic CPS framework considers m different types of events (shocks) which may occur several times. A Poisson process $N_t^{(e)}$ with intensity $\lambda^{(e)}$ counts the number of occurrences of the e event, that is, the process

$$\{N_t^{(e)}, \quad t \geq 0\} \tag{5.42}$$

records the number of times the event of type e has occurred in $(0, t]$. We can think of these events as being exogenous shocks as well as endogenous perturbations of the market. The CPS model assumes the event counting processes for different events to be independent. Now suppose there are n different asset prices and that the r^{th} occurrence of event of type e triggers price j to jump with a certain probability $p_{r,j}^{(e)}$. This is formalised by the process

$$\{N_{j,t}, \quad t \geq 0\} \tag{5.43}$$

with

$$N_{j,t} = \sum_{e=1}^{m} \sum_{r=1}^{N_t^{(e)}} I_{j,r}^{(e)}. \tag{5.44}$$

where $I_{j,r}^{(e)}$ is a Bernoulli variable with probability $\mathbb{P}(I_{j,r}^{(e)} = 1) = p_{r,j}^{(e)}$. The two assumptions of Poisson process for $N_t^{(e)}$ and of Bernoulli process for $I_{j,r}^{(e)}$ make the process in equation (5.44) Poisson itself. The core aspect to note is that same events trigger jumps for different assets and thus introduce a dependence between them. In other words, for a fixed event the variables which trigger jumps of different assets may be dependent. The dependence structure is encoded in the multivariate Bernoulli distribution. Consider the vector of Bernoulli variables

$$\mathbf{I}_r^{(e)} = (I_{1,r}^{(e)}, \ldots, I_{n,r}^{(e)}). \tag{5.45}$$

The joint probability distribution of having jumps in n different stock prices (dropping the r to simplify the notation) is given by

$$\mathbb{P}(I_1^{(e)} = i_1, \ldots, I_n^{(e)} = i_n) = p_{1,\ldots,n}^{(e)}(i_1, \ldots, i_n), \quad i_1, \ldots, i_n \in \{0, 1\}. \tag{5.46}$$

In the case of conditional independence this reads

$$p_{1,\ldots,n}^{(e)}(1, \ldots, 1) = \prod_{j=1}^{n} p_j^{(e)}, \tag{5.47}$$

where $p_j^{(e)} = p_j^{(e)}(1)$ are the one dimensional marginal probabilities. When

$$p_{1,\ldots,n}^{(e)}(1, \ldots, 1) \geq \prod_{j=1}^{n} p_j^{(e)} \tag{5.48}$$

the jumps are said to be positively dependent, with the strongest possible dependence given by

$$p_{1,\ldots,n}^{(e)}(1, \ldots, 1) = \min_j(\, p_1^{(e)}, \ldots, p_n^{(e)}). \tag{5.49}$$

In order to render this multivariate Poisson model realistic in the description of stock returns, a further random variable describing the amplitudes of the jumps must be introduced. The Poisson process now becomes a *compound* Poisson process. Whether the process actually jumps or not is still determined by the indicator function $I_{j,r}^{(e)}$, but the jump size distribution may depend on the event type e. The new process is

$$Z_{j,t} = \sum_{e=1}^{m} \sum_{r=1}^{N_t^{(e)}} I_{j,r}^{(e)} X_{j,r}^{(e)} \tag{5.50}$$

where the vector $\mathbf{X}_{j,r}^{(e)}$ is distributed according to the multivariate distribution $F^{(e)}$. Note that the $F^{(e)}$ depends on the event type e but not on j, hence the amplitude of the jump has the same law for all processes $\{Z_j : j = 1, \ldots, n\}$. This means that the activity of the market as measured by the fluctuation amplitudes is not a property of the assets, rather of the type of shocks the assets are exposed to. The resulting total univariate process for a given j is thus non-stationary, but rather

an aggregate of stochastic contributions coming from several distributions.
By taking the expectation value of equation (5.44) and using tower property for the conditional expectation we get the total intensity λ_{N_j} of the Poisson process $N_{j,t}$:

$$\begin{aligned} E\left(N_{j,t}\right) &= E\left(\sum_{e=1}^{m}\sum_{r=1}^{N_t^{(e)}} I_{j,r}^{(e)}\right) = \sum_{e=1}^{m} E\left[E\left(\sum_{r=1}^{n_t^{(e)}} I_{j,r}^{(e)} \;\middle|\; N_t^{(e)}\right)\right] \\ &= \sum_{e=1}^{m} E\left[N_t^{(e)} E\left(I_{j,r}^{(e)}\right)\right] \\ &= t\sum_{e=1}^{m} \lambda^{(e)} p_j^{(e)} = \lambda_{N_j}. \end{aligned}$$

This clearly shows that for a fixed event type e, the intensity of the jump process for asset j is shrunk according to how likely it is that the events $\{e = 1, \dots k,\ k < m\}$ are relevant for that asset. The Bernoulli variable $I_{j,r}^{(e)}$ has thus the effect of conditioning further the jump of j, given that the event e has occurred. This condition allows to weight the impact of common shocks on a basket of stocks and hence provides a straightforward way to organise the market in sectors according to different types of risk. Note also that different events have different intensities and therefore a hierarchy of rare events can be established. In the next section we elaborate on the basis of this formalism to build a mechanism of dependence between asset prices.

5.5.2 Common Hawkes Shocks

To fully exploit the formalism introduced in Section 5.5.1 to construct a multivariate stochastic volatility model, one needs to introduce a temporal dependency structure in the underlying point process and to this end, Poisson processes are not suitable. A possible way to achieve the desired temporal structure is through Hawkes processes, first introduced in [25]. Hawkes processes are a particular type of linear self-excited point processes [169]. A one dimensional Hakwes process is

a counting process N_t whose intensity reads [25]:

$$\begin{aligned}\lambda_t &= \lambda_\infty + \int_{-\infty}^{t} \nu_{t-s} dN_s \\ &= \lambda_\infty + \sum_{u:t_u<t} \nu_{t-t_u}\end{aligned} \tag{5.51}$$

where λ_∞ is a constant intensity and $\nu\colon \mathbb{R}^+ \mapsto \mathbb{R}^+$ is a kernel expressing the influence of past events t_u on the current intensity process λ_t, whose analytical form can be specified. For their capability of incorporating a feedback effect via the self-excited intensity λ_t, the employment of Hawkes processes spans across several scientific disciplines, from seismology [170, 171] to signal processing and financial markets [172, 173]. In the latter context their popularity has been growing more and more over the last few years, with particular attention to market microstructure mechanism [174–176], but also to modelling contagion in daily stock returns [177].

Analogously to compound Poisson processes one can associate with the arrival of a Hawkes process a fluctuation size through a random variable $\mathbf{X}$ with a corresponding distribution function $F_{\mathbf{X}}(\mathbf{x})$.

The kernel originally specified by Hawkes in [25] and most often considered in the literature is the exponential kernel

$$\nu_t = \alpha e^{-\beta t}. \tag{5.52}$$

Chavez-Demoulin *et al.* [173] consider, following [170], a kernel that, besides the dependence on past events, depends also on the size of previous fluctuations: if one labels by x_u the size of the fluctuation at time t_u then the kernel reads

$$\nu_{t-t_u;x_u} = \frac{\psi e^{\beta x_u}}{(t-t_u+\gamma)^{\rho+1}}, \qquad t > t_u \tag{5.53}$$

where $\psi, \beta, \rho, \gamma$ are strictly positive constants. With this specification the intensity of present shocks also critically depends (exponentially) on the amplitude of the fluctuations recorded in the past.

We construct a multivariate stochastic volatility model, that we dub Common

Hawkes Shock model (CHS), considering m different Hawkes processes $\{N_t^{(e)} : e = 1, \dots, m\}$, each one associated with a risk factor e and with exponential intensity as in Equation (5.52). These processes should be interpreted as the risks affecting each stock for a given time t. Each risk has a proper intensity, which is specified by the set of parameters $\{\lambda_\infty^{(e)}, \alpha^{(e)}, \beta^{(e)}\}$. With reference to the general factor multiplicative models of Section 5.4 we model returns as

$$r_{i,t} = \exp(z_{i,t})\epsilon_{i,t} \tag{5.54}$$

with, as usual, $\boldsymbol{\epsilon}_t = (\epsilon_{1,t}, \dots, \epsilon_{N,t})$ a multivariate Gaussian vector $\boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \Sigma_\epsilon)$ and

$$z_{i,t} = \sum_{e=1}^{m} \sum_{r=1}^{N_t^{(e)}} I_{i,r}^{(e)} X_{i,r}^{(e)} \tag{5.55}$$

where $N_t^{(e)}$ are now Hawkes processes. The random variables $X_{i,r}^{(e)}$ are, for each $i = 1, \dots, N$ and for each $e = 1, \dots, m$, the size of the fluctuations in the stock i due to risk of type e. For each time step t, each risk e strikes $N_t^{(e)}$ times. We set the random vector $\mathbf{X}^{(e)}$ for all risks $e = 1, \dots, m$ to have independent components, which results in the multivariate distributions $F_{\mathbf{X}}^{(e)}(\mathbf{x})$ to be factorized as

$$F_{\mathbf{X}}^{(e)}(\mathbf{x}) = F_{X_1}^{(e)}(x_1) \cdot \dots \cdot F_{X_N}^{(e)}(x_N). \tag{5.56}$$

The random variable in Equation (5.55) is thus a multifactor additive log-volatility which, thanks to the exponential function, gives a multifactor multiplicative volatility so that returns can be written as

$$r_{i,t} = \prod_{e=1}^{m} \exp\left(\sum_{r=1}^{N_t^{(e)}} I_{i,r}^{(e)} X_{i,r}^{(e)} \right) \epsilon_{i,t}. \tag{5.57}$$

In the specification of Equation (5.57) different risks affect different groups of stocks, as schematically represented in Figure 5.5. In the example reported in the figure, there is one risk common to all stocks ($\lambda^{(4)}$) which can be thought of as the only common risk in the one-factor multiplicative model of Equation (5.33) and then other risks affect groups of stocks locally. The CHS can therefore be seen as

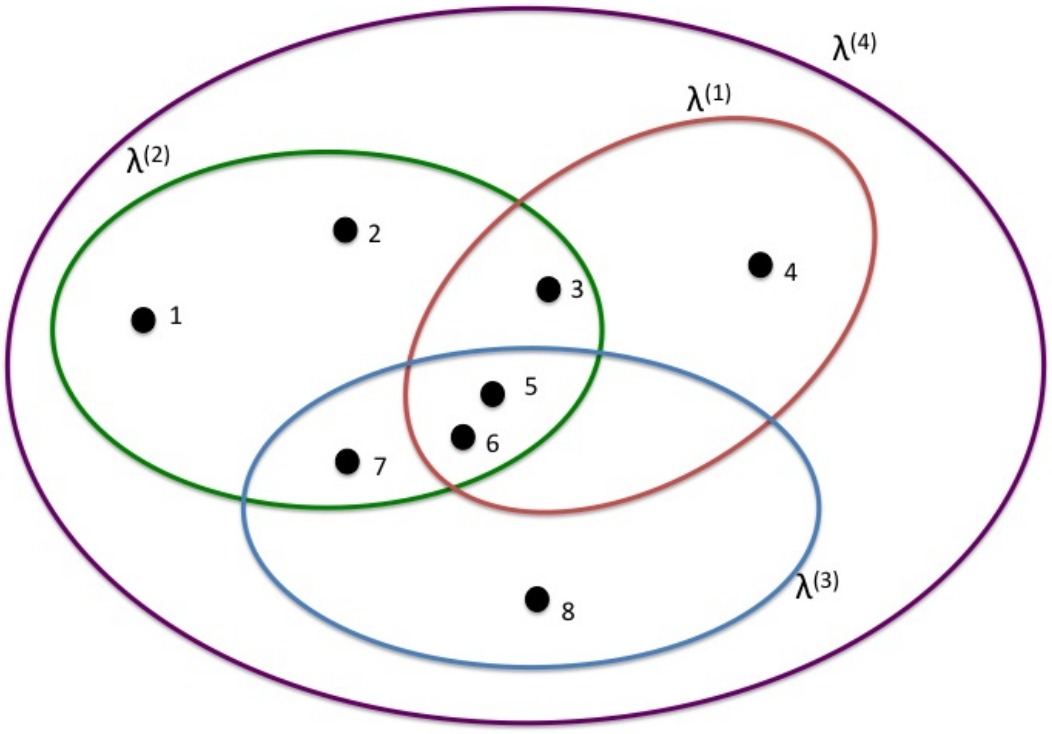

FIGURE 5.5: A schematic representation of the hierarchical influence of different risks with corresponding intensity $\lambda^{(e)}$, for $e = 1, \ldots, 4$. The black dots are stocks and the various ovals represent different risks.

a generalisation of one-factor multiplicative models, with the further feature that the volatility bursts are now conditional on the arrival of Hawkes processes. To make an explicit example, let us write explicitly the dynamics of stock labelled by 3 in Figure 5.5. This stock is affected by three different risks, namely $\lambda^{(4)}, \lambda^{(1)}$ and $\lambda^{(2)}$. Equation (5.57) then turns into

$$r_{3,t} = \exp\left(\sum_{r=1}^{N_t^{(1)}} I_{3,r}^{(1)} X_{3,r}^{(1)} + \sum_{r=1}^{N_t^{(2)}} I_{3,r}^{(2)} X_{3,r}^{(2)} + \sum_{r=1}^{N_t^{(4)}} I_{3,r}^{(4)} X_{3,r}^{(4)}\right) \epsilon_{3,t}. \tag{5.58}$$

The actual contribution of each of the three sums in the exponential will depend on both (i) the values of the intensities $\lambda_t^{(e)}$ (ii) the probabilities associated to each risk type $p_i^{(e)}$. Very low intensities will let the underlying Hawkes process jump rarely and, conditional on the jump, the contribution of the variable $X^{(e)}$ to the overall volatility depends on the value of the Bernoulli probability $p^{(e)}$: a

value of $p^{(e)}$ close to unity is likely to make the corresponding underlying Hawkes process effective on the volatility.

Let us give a close look at the covariance structure of the log-volatility process z_t (we drop the label i as we concentrate on one arbitrary time series). We have

$$\mathrm{Cov}[dz_t, dz_{t+\tau}] = \mathbb{E}[dz_t dz_{t+\tau}] - \mathbb{E}[dz_t]\mathbb{E}[dz_{t+\tau}], \tag{5.59}$$

where

$$dz_t = \sum_{e=1}^{m} I_t^{(e)} X_t^{(e)} dN_t^{(e)}. \tag{5.60}$$

From this expression the second member in the right hand side of equation (5.59) is readily computed by considering all three r.v.'s X, I and N independent from each other:

$$\begin{aligned}
\mathbb{E}[dz_t] &= \mathbb{E}\left[\sum_{e=1}^{m} I^{(r)} X^{(e)} dN_t^{(e)}\right] \\
&= \sum_{e=1}^{m} \mathbb{E}[I^{(e)}]\, \mathbb{E}[X_t^{(e)}]\, \mathbb{E}[dN_t^{(e)}] \\
&= \sum_{e=1}^{m} p_t^{(e)}\, \mathbb{E}[X_t^{(e)}]\, \lambda^{(e)} dt,
\end{aligned} \tag{5.61}$$

where $\lambda^{(e)} = \mathbb{E}[\lambda_t^{(e)}]$. Analogously we have

$$\mathbb{E}[dz_{t+\tau}] = \sum_{e=1}^{m} p_{t+\tau}^{(e)}\, \mathbb{E}[X_{t+\tau}]^{(e)}]\, \lambda^{(e)}\, dt. \tag{5.62}$$

For the mixed term $\mathbb{E}[dz_t dz_{t+\tau}]$ we have instead

$$\mathbb{E}[dz_t dz_{t+\tau}] = \sum_{e=1}^{m} \mathbb{E}[X_t^{(e)} X_{t+\tau}^{(e)}]\, \mathbb{E}[I_t^{(e)} I_{t+\tau}^{(e)}]\, \mathbb{E}[dN_t^{(e)} dN_{t+\tau}^{(e)}]. \tag{5.63}$$

For an exponential kernel of the form (5.52), we exploit the result found by Hawkes in [25], that is

$$(\mathbb{E}[dN_t dN_{t+\tau}] - \lambda^2)/dt^2 = \frac{\beta\, \lambda\, (2\alpha - \beta)}{2(\alpha - \beta)} e^{-(\alpha-\beta)\,\tau}. \tag{5.64}$$

We obtain

$$\mathbb{E}[dN_t^{(e)} dN_{t+\tau}^{(e)}] = (g^{(e)}(\tau) + {\lambda^{(e)}}^2)\, dt^2 \tag{5.65}$$

where

$$g(\tau) = \frac{\beta\,\lambda\,(2\alpha - \beta)}{2(\alpha - \beta)} e^{-(\alpha-\beta)\,\tau} \tag{5.66}$$

and the $g^{(e)}(\tau)$ is obtained by giving to all parameters in equation (5.66) an e-dependence. Finally, putting all pieces together and using the fact that, because of stationarity of the $X^{(e)}$

$$\mathbb{E}[X_t^{(e)}] = \mathbb{E}[X_{t+\tau}^{(e)}] = \mathbb{E}[X^{(e)}], \tag{5.67}$$

together with the independence of the $X^{(e)}$ and of the $I^{(e)}$ we have

$$\mathrm{Cov}[dz_t, dz_{t+\tau}] = \sum_{e=1}^{m} \mathbb{E}[X^{(e)}]^2\, p_t^{(e)}\, p_{t+\tau}^{(e)}\, \left(g^{(e)}(\tau) - (\lambda^{(e)})^2\right) dt^2. \tag{5.68}$$

Since one further assumes the probability of reaction to risk e to be constant in time, we finally have

$$\mathrm{Cov}[dz_t, dz_{t+\tau}] = \sum_{e=1}^{m} \mathbb{E}[X^{(e)}]^2\, (p_t^{(e)})^2\, \left(g^{(e)}(\tau) - (\lambda^{(e)})^2\right) dt^2. \tag{5.69}$$

The log-volatility covariance thus results in a superposition of exponential decays, with m different integral scales

$$T^{(e)} = \frac{1}{\alpha^{(e)} - \beta^{(e)}} \tag{5.70}$$

and each integral scale $T^{(e)}$ is associated with a specific risk e. The presence of heterogeneous time scales in the volatility auto-correlation has been already proposed in several stochastic models [89, 90, 178]. In [89] the authors discuss how one single exponential time scale, as proposed for example in stochastic volatility models like [179] or in the GARCH specification, cannot explain the long-ranged dependence observed in the volatility. In fact, if the time scale T is chosen to be very small (of the order of few days) then the long tail is completely missed, whereas if the time scale is of the order of hundreds of days, one misses the

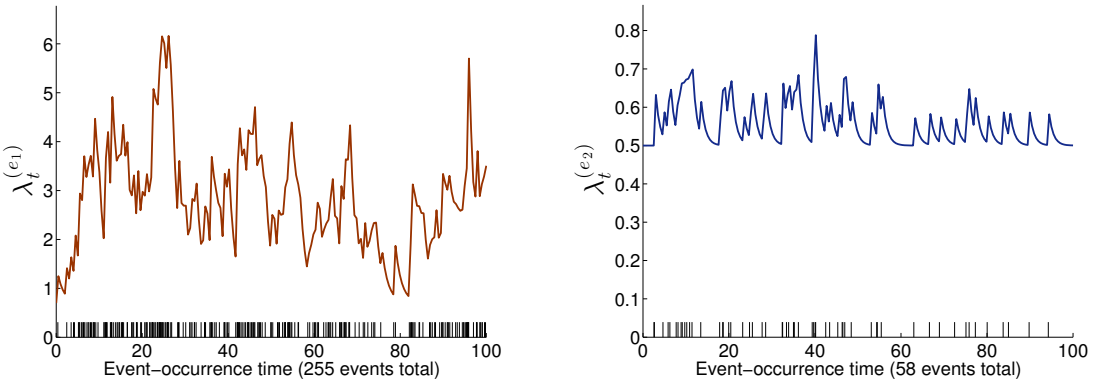

FIGURE 5.6: Simulated intensity processes $\lambda^{(e_1)}$ and $\lambda^{(e_2)}$ for two risks e_1 and e_2. The parameters of the Hawkes intensity λ_t are set to $\alpha = 0.6$, $\beta = 0.7$ and $\lambda_\infty = 0.7$ for e_1 and $\alpha = 0.1$, $\beta = 1$ and $\lambda_\infty = 0.5$ for e_2.

rapid decrease for small τ. One needs at least *two* time scales in order to mimic the behaviour shown by the data. In the CHS setting the appearance of many exponential time scales has a straightforward interpretation as the heterogeneous persistence of different risks upon stock returns.

In Figure 5.6 we show two realisations of the intensities of the two underlying risks $\lambda^{(e_1)}$ and $\lambda^{(e_2)}$, with different values of parameters α, β and λ_∞. The plots also show the occurrence times of the events which triggers bursts in the volatility. The different values of the parameter set $\{\alpha, \beta, \lambda_\infty\}$ implicitly establishes a hierarchy in the frequency of risks, as can be appreciated from the very different total number of events for the two risks. It is therefore natural in this framework to associate risks of different type and frequency with the $\lambda^{(e)}$.

The multivariate model explored in this paragraph, although appealing for its intuitive ground, is hardly tractable because of the $\lambda^{(e)}$ being unobservable and because of the large number of parameters needed. In particular, one needs to specify the intensities of all the underlying risks together with the probabilities of reaction of each stock to each risk: these are $m(N+1)$ extra-parameters, let alone the parameters of the distributions $F^{(e)}$, which are in principle arbitrary. We thought it nonetheless worthy of discussion as a first step towards a multivariate model incorporating a hierarchy of the risks. In the last chapter of this thesis we are going to see how the explicit inclusion of a hierarchical structure of risks

into a multivariate model is able to explain dependency patterns observed between hierarchical cross-correlation structure and multifractal properties.

5.6 Conclusions

In this chapter we have reviewed multivariate models used to describe the dependency structure of asset returns. We have first discussed static elliptical models, remarking how they lack any reference to the temporal properties observed in empirical data, yet providing a faithful reference framework for further modelling techniques. We have then explained how both MGARCH and MSV models can succeed in describing the time-dependence together with the cross-sectional dependence, although one necessarily ends up having more complex and sometimes over-parametric models. We have further discussed how factor models can simplify these models by reducing the number of relevant factors needed to describe the dynamics of financial time series and eventually we have investigated a new multi factor volatility model based on Common Shock formalism. We have pointed out the limits of this approach, yet remarking how it provides a possible way of introducing a hierarchical structure of risk factors in a multivariate model, a feature which most of the models present in the literature are lacking.

Chapter 6

A Multivariate Dynamical Hierarchical Model

"There's no sense in being precise when you don't even know what you are talking about."

- J. Von Neumann -

In this chapter we introduce an alternative model to multivariate stochastic volatility models [136] discussed in Chapter 5. Differently from all models in the literature this model includes explicitly a time varying hierarchical structure of risks, which is shown to be able to reproduce dependence between multifractality and hierarchical order, the new empirical fact which has been discussed in Chapter 4. The effect of the parameters of the model on the scaling as well as cross-correlation properties are discussed in details and comparison, in particular, with elliptical models [23] and empirical data are also presented. The material presented in this chapter is part of the paper "Dependency Structure and Scaling Properties of Financial Time Series are Related", published Scientific Reports [21].

6.1 Introduction

In this chapter we propose a new all-encompassing mechanism able to account coherently both for time-varying multifractality and dependence between scaling properties and hierarchical structure. Since these two empirical properties we want to describe have never been investigated before, there are, to the best of our knowledge, no models attempting to reproduce these features standing as reference in the literature. Starting from multivariate stochastic volatility models discussed in Chapter 5, we introduce a new mechanism that enables us to reproduce the empirical facts found out in the previous chapters of this thesis. The novelty of our model lies in the introduction of a hierarchical term, factorized from a common volatility factor. This construction adds richness to standard multivariate volatility models incorporating explicitly the market hierarchical structure and allows to recover MSV models in certain regimes.

Although, as we have discussed in Chapter 4, the network tools technique have been proved to be successful in describing the dependency structure in financial markets, only very few studies can be retrieved in the literature trying to associate the cross-correlation hierarchicy to the factor structure and incorporate it into a multivariate model. The most relevant one is the Hierarchically Nested Factor Model (HNFM) introduced by Tumminello *et al.* in [110]. In this work the authors show that a multivariate factor model can be straightforwardly associated with a ALCA in such a way that the correlation matrix of the factor model is the one which produce the hierarchical tree of the ALCA. To obtain this characterisation the authors introduce a set of hierarchical factors whose weights are directly related to the filtered correlation matrix associated with the hierarchical tree obtained via ALCA. The HNFM provides a first framework to incorporate the hierarchical structure into a multivariate model, but lacks any type of detail about the temporal structure of the factors which is of paramount importance in the description of financial time series dynamics.

More recently, Chicheportiche and Bouchaud, starting from the observation that

the joint stock returns distributions systematically defy the hypothesis of elliptical structure [142], have proposed a nested factor model that accounts for all relevant non-linearities in the dependency structure [180]. However, similarly to the HNFM, this model doesn't take into account the temporal structure of the factors, that is instead presented in our formulation. The model proposed in this chapter represent a first attempt to study the interplay between cross-correlation hierarchical structure and multifractality.

Throughout this chapter we use the dataset of NYSE daily closing prices for 342 most capitalised companies in the period 1997-2012, also used in Chapter 3. The data are used both as comparison with the model and to extract the hierarchical structure needed to simulate the model.

This chapter is organised as follows: in Section 6.2 we introduce the model and describe its properties also in relation with elliptical models. In Section 6.3 we discuss in detail how the multifractal properties of the model crucially depend on the hierarchical structure and how time-varying multifractality can be recovered by switching between different hierarchical regimes. In this same section we also show the ability of the model to reproduce multifractality positively correlated with correlation hierarchical order. Conclusions are given in Section 6.4.

6.2 Model Construction and Properties

In the spirit of generalising log-normal multivariate models with one single volatility factor [143], we model the volatility of stock i as

$$\sigma_{i,t} = w_t \prod_{m \in \gamma_{i,t}} e^{K_m} \tag{6.1}$$

where K_m are Bernoulli random variables with probabilities p_m associated to the nodes a_m along the hierarchical tree $\Gamma_{i,t}$, that is

$$K_m = \begin{cases} 1, & p_m \\ 0, & 1 - p_m. \end{cases} \tag{6.2}$$

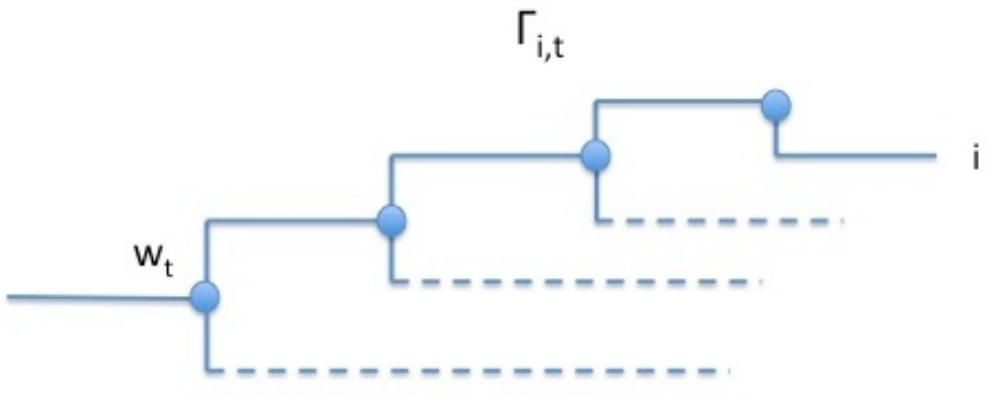

FIGURE 6.1: The hierarchical tree $\Gamma_{i,t} = \{a_m,\ m \in \gamma_{i,t}\}$ associated with stock i. The volatility factor w_t is common to all stocks.

Note that, in this way, the volatility $\sigma_{i,t}$ is let explicitly depend on the hierarchical path γ_i (or equivalently on the set of nodes $\Gamma_{i,t} = \{a_m,\ m \in \gamma_{i,t}\}$), as shown in Figure 6.1. The process w_t is chosen as a stationary log-normal stochastic process autocorrelated in time, with the autocovariance function decaying as a power law, i.e. $\mathrm{Cov}(w_t w_{t+\ell}) \sim \ell^{-1}$, in line with [20]. Returns therefore take the form

$$r_{i,t} = \sigma_{i,t}\,\epsilon_{i,t} = Y^{(i)}_{n,t}\,w_t\,\epsilon_{i,t} \tag{6.3}$$

where $Y^{(i)}_{n,t} = \prod_{m\in\gamma_{i,t}} e^{K_m}$. We call this construction of Equations (6.1) and (6.3) Multivariate Dynamical Hierarchical Model (MDHM or simply DHM). It is worth remarking that the hierarchical term $Y^{(i)}_{n,t}$ introduces a richer structure of dependence, where the topology of the risk organisation plays a crucial role in creating heterogeneous dependence not accounted for by standard multivariate models. The effect of the hierarchy can be seen as a perturbation to a dominant volatility which is responsible for the long-range dependence observed in stock returns.

As schematically sketched in Figure 6.2, the hierarchical structure can be assumed to be time-varying: each stock is affected by different risks in time and shares these risks with diverse groups of other stocks, the topology of the hierarchy necessarily evolving in time to reflect the changing market complexity.

One can think of each node in the hierarchical structure as a particular trigger for selected stocks. For a specific risk m the variable K_m represents its impact on

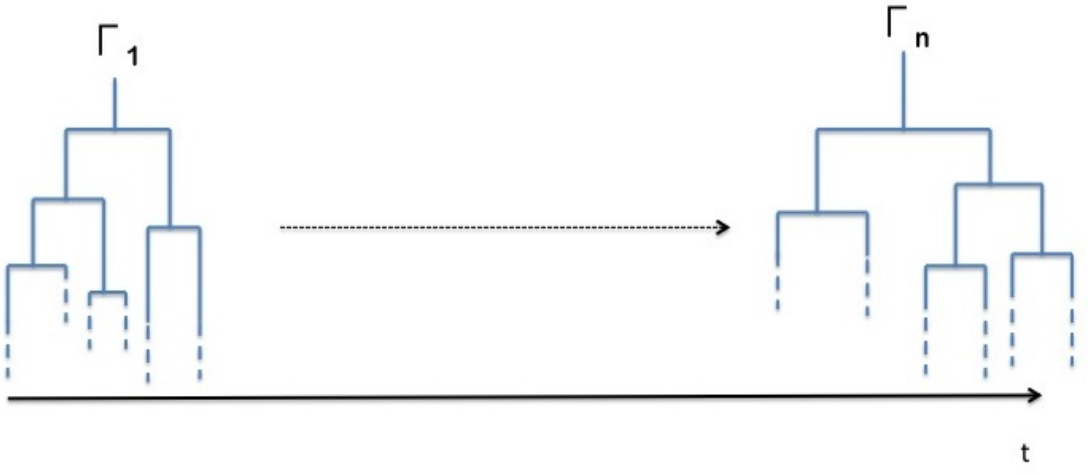

FIGURE 6.2: We show a schematic example of the time evolution of the cross-correlation hierarchical structure, with the hierarchical tree Γ_t changing from Γ_1 at time $t = 1$ to Γ_n at $t = n$.

the market. The parameters p_m are thus crucial in determining the complexity of the web of risks. Values of $p_m \sim 1$ characterise a risk that is very likely to affect the stocks exposed to it. Conversely, $p_m \sim 0$ implies that risk m has a very low incidence: the dynamics of stocks exposed to it is unlikely to actually be affected. Note that the amplitude of the fluctuations triggered by risk m is totally uncorrelated from the value of the corresponding p_m, which means that a very rare risk can still trigger large fluctuations if combined with the effect of others. In this thesis we have considered the simplest hypothesis of risks uncorrelated between each other, with each probability p_m assumed to be independent from the others and randomly drawn in specific sub-intervals in $[0, 1]$.

In the following subsection we investigate the relationship between our model and MSV models.

6.2.1 Comparison with elliptical models

Comparing MVS of Section 5.3.2 and particularly Section 5.4.2 with Equation (6.3) one can see that the MDHM generalises MSV models, which are recovered in the limit $Y_{n,t} \to 1$ in Equation (6.3). Note that the presence of a non-trivial time

correlation in the volatility w_t doesn't spoil the elliptical structure, provided w_t is a log-normal process: the MSV obtained for $Y_{n,t} = 1$ is therefore still elliptical.
In the MDHM, the volatility factor $\sigma_{i,t}$ is the same only for those stocks sharing the same hierarchical structure of risks, a feature not present in MSV models where the volatility is the same for all stocks. Therefore the explicit introduction of a hierarchical structure in modelling the dynamics gives the multivariate model a richer structure. Only those assets sharing exactly the same hierarchical structure share the same multivariate structure.
The inadequacy of elliptical model to describe the dependency structure in stock markets has been recently reported in [142]. In particular, the authors have shown that weakly correlated stocks are the ones found most often in disagreement with the elliptical hypothesis. The intuition behind their finding was that one single volatility factor may not be enough to describe the complexity of the market structure.
In Figure 6.3 we plot linear correlation coefficient against Kendall's τ for log-normal multivariate model (yellow dots), DHM (green dots) and empirical data (purple dots). We also show in red thick line the theoretical relation expected to hold between correlation and Kendall τ for multivariate elliptical random variables, which was established in [181] to be

$$\tau = \frac{2}{\pi} \arcsin \rho. \tag{6.4}$$

We can see how DHM departs from the elliptical structure, especially in the low correlation regime, whereas the log-normal model is, by construction, faithful to the relation (6.4). Note that, among the green dots, some preserve the elliptical structure: these are those processes sharing the same hierarchy and thus not affected by the factor $Y_{n,t}$. The DHM thus predicts a finer structure of risks factors affecting the stocks heterogeneously, in such a way that a simple multivariate model cannot account for. We must remark though that the empirical findings reported in [142] have led the authors to rule out any pseudo-elliptical generalisations of the form $r_i = \sigma_i \epsilon_i$ as possible models for non-linearities in the joint stock

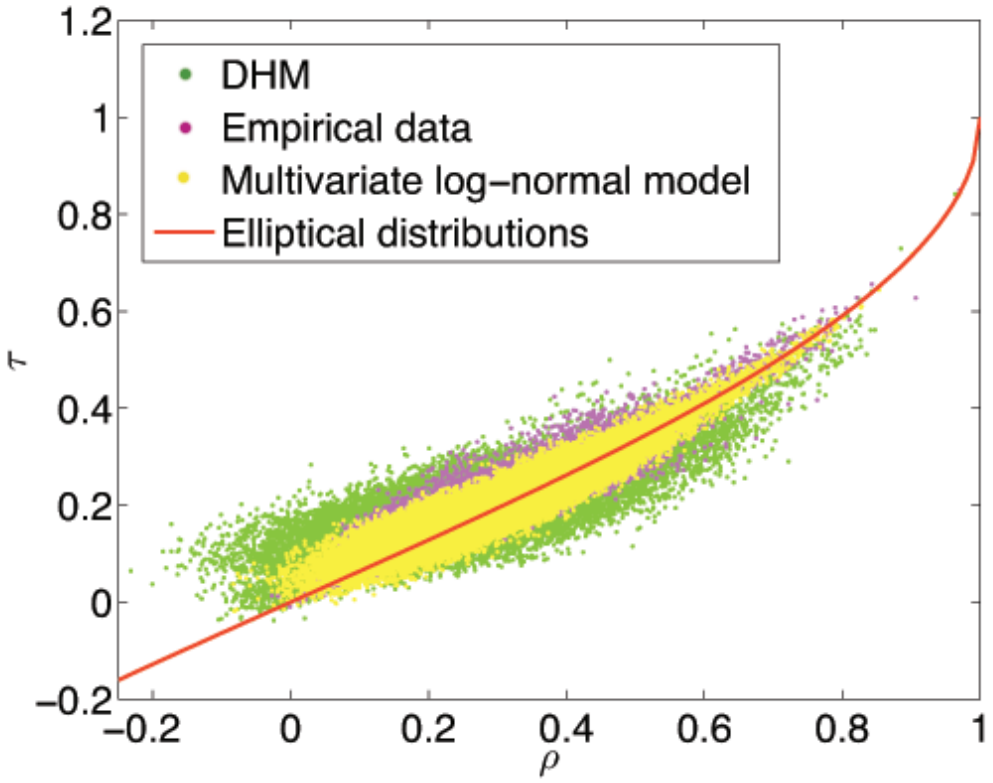

FIGURE 6.3: We plot Kendall τ against linear correlation ρ for multivariate log-normally distributed synthetic time series (yellow dots), synthetic DHM time series (green dots) and empirical data (magenta dots), compared with the theoretical relation expected for elliptically distributed random variables (red thick line). The DHM has been simulated with probabilities drawn randomly in the interval $[0.4, 0.6]$.

returns distribution. Such models are in fact incapable to account for specific observed properties of the copula. Therefore, in order to better include also a finer dependency characterisation, the MDHM needs to be somehow tweaked according to the findings of Chicheportiche and Bouchaud [142, 143, 180] .

In the next subsections we are going to give some analytical details about the distribution characterising the DHM volatility and the associated moments. We will then derive exact expressions for the cross-correlation coefficients and the volatility autocorrelation function.

6.2.2 Distribution of the Volatility Process

We want to find a close form expression for the distribution of the process

$$Y_{n,t}^{(i)} = \prod_{m \in \gamma_{i,t}} e^{K_m} = \prod_{m \in \gamma_{i,t}} \tilde{K}_m. \tag{6.5}$$

We drop the time indices t and i to simplify the notation. At each time step t the distribution of each K_m random variable is stationary and reads

$$p\left(k_m\right) = \binom{1}{k_m} p_m^{k_m} (1-p_m)^{1-k_m}, \tag{6.6}$$

where $k_m = \{0, 1\}$ are the values that can be attained by the random variables K_m. It follows that each $\tilde{K}_m$ can assume the two values $\{1, e\}$ with probabilities $1 - p_m$ and p_m respectively. In order to write the distribution in a compact form, we will make use of the following notation:

Definition 3. Let m be a positive integer. For $\ell = 0, \ldots, m$, $\mathcal{P}_m^\ell$ is the set of ℓ-tuples of numbers $1, \ldots, m$, that is

$$\begin{cases} \mathcal{P}_m^0 = \{\emptyset\} \\ \mathcal{P}_m^1 = \{\{1\}, \{2\}, \{3\}, \ldots, \{m\}\} \\ \mathcal{P}_m^2 = \{\{1,2\}, \{1,3\}, \{1,4\}, \ldots, \{1,m\}, \{2,3\}, \{2,4\}, \ldots, \{2,m\}, \ldots, \{m-1,m\}\} \\ \vdots \\ \mathcal{P}_m^m = \{\{1,2,3,\ldots,m\}\}. \end{cases}$$

Proposition 1. The random variable Y_n assumes values in $\{e^0, e^1, e^2, \ldots, e^n\}$ and the probability of assuming the general value e^ℓ is

$$p(Y_n = e^\ell) = \sum_{\nu \in \mathcal{P}_n^\ell} \prod_{\nu, \mu \neq \nu} p_\nu (1 - p_\mu), \qquad \mu = 1, \ldots, n. \tag{6.7}$$

Proof. Y_n is the product of n $\tilde{K}_m$, each of which can be either 1 or e. Therefore the possible values for Y_n are those given in Table 6.1.

Note that, apart from the two extreme cases where $\tilde{K}_1 \ldots \tilde{K}_n$ equal all 1 or e, there

$\tilde{K}_1$	$\tilde{K}_2$	$\tilde{K}_3$	$\dots$	$\tilde{K}_n$	Y_n
1	1	1	$\dots$	1	1
e	1	1	$\dots$	1	e
1	e	1	$\dots$	1	e
1	1	e	$\dots$	1	e
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
1	1	1	$\dots$	e	e
e	e	1	$\dots$	1	e^2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
e	e	e	$\dots$	1	e^3
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
e	e	e	$\dots$	e	e^n.

TABLE 6.1: The possible values of Y_n corresponding to all combinations of $\tilde{K}_m,\ m = 1, \dots, n$.

are in general different combinations of the $\tilde{K}$ which can yield the same power of e. The probability of each of the values in the last column is obtained summing the products of all possible different combinations of $\tilde{K}_m$, $m = 1, \dots, n$, yielding the corresponding Y_n. Specifically we have

$$\begin{cases} p(Y_n = e^0) = (1 - p_1) \times \cdots \times (1 - p_n) = \prod_{i=1}^{n} (1 - p_i) \\ p(Y_n = e^1) = p_1 \times \prod_{i=2}^{n} (1 - p_i) + (1 - p_1)\, p_2 \times \prod_{i=3}^{n} (1 - p_i) + \dots \\ + \prod_{i=1}^{n-1} (1 - p_i) \times p_n \\ \vdots \\ p(Y_n = e^n) = \prod_{i=1}^{n} p_i. \end{cases}$$

Let us concentrate, without loss of generality, on the expression for $p(Y_n = e^1)$. Exploiting the notation introduced in Definition 3, we can re-write $p(Y_n = e^1)$ as

$$p(Y_n = e^1) = \sum_{\nu \in \mathcal{P}_n^1} \prod_{\nu, \mu \neq \nu} p_\nu (1 - p_\mu)$$

where $\mu = 1, \ldots, n$ and the values of μ in the ℓ-tuple ν are excluded from the product. Analogously, we have

$$p(Y_n = e^2) = \sum_{\nu \in \mathcal{P}_n^2} \prod_{\nu, \mu \neq \nu} p_\nu (1 - p_\mu)$$

and, for the general ℓ

$$p(Y_n = e^\ell) = \sum_{\nu \in \mathcal{P}_n^\ell} \prod_{\nu, \mu \neq \nu} p_\nu (1 - p_\mu).$$

□

Definition 4. We shall refer to the probability density function of the random variable $Y_{n,t}$ given in equations (6.7) as $f(y)$, with y taking values in $\{e^0, e^1, e^2, \ldots, e^n\}$.

Let us give a closer look at the properties of this distribution. We plot in Figure 6.4 the probability mass function $p(Y_n = e^\ell)$ obtained from simulations of the process Y_n of length 2000 for different values of n and different values of the parameters p's. For visualisation purposes we plot the histogram in semi-log scale. The mass function is symmetric for $p_m = 0.5,\ m = 1, \ldots, n$ but this symmetry is lost as soon as the parameter values are changed. Specifically, $p_m > 0.5$ shifts the bulk toward large values, whereas $p_m < 0.5$ shifts the bulk toward small values. This reflects the intuitive fact that the contribution of the process $Y_{n,t}$ is more relevant when the probabilities of the Bernoulli variables are large. As a result, the impact of the hierarchical risks onto the distribution tails is proportional to the probability of the relevant risks being effective.

Next, we study the distribution of the overall volatility of stock i, namely the process

$$\sigma_t = Y_{n,t} w_t. \tag{6.8}$$

We take w_t to be a log-normal random variable, with density

$$g(w) = \frac{1}{\sqrt{2\pi} s w} \exp\left(-\frac{(\log w)^2}{2s^2}\right), \qquad w > 0, \tag{6.9}$$

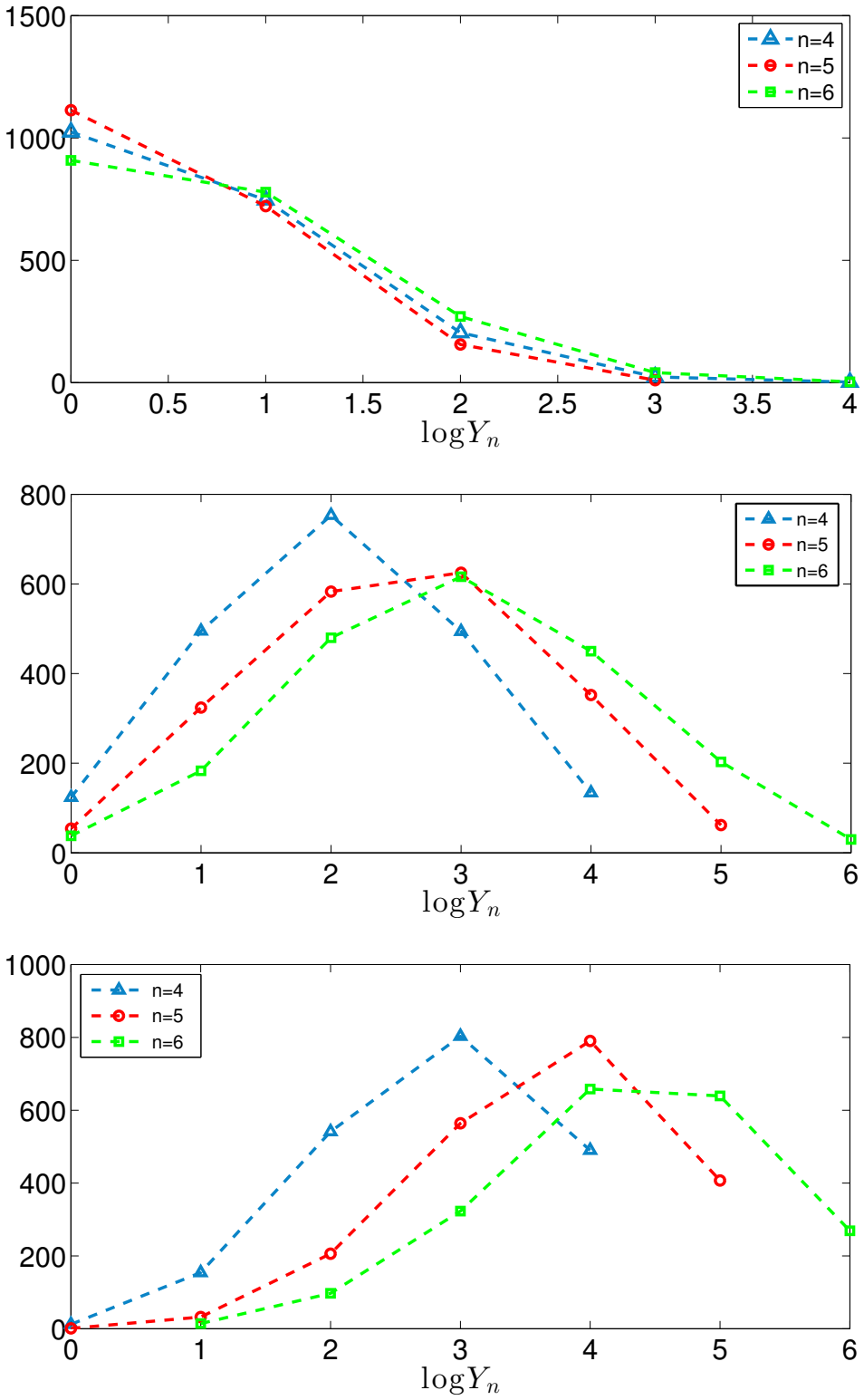

FIGURE 6.4: We plot the histograms of $\log Y_n$ obtained from simulations of the process Y_n of length 2000 for several values of n and in different ranges of the probabilities: $p \in [0, 0.2]$ (top), $p = 0.5$ (middle) and $p \in [0.6, 0.8]$ (bottom).

where s is the shape parameter. The variance of the random variable w_t is $v_w = (e^{s^2} - 1)e^{s^2}$ [48].

The probability density function of σ_t, which we shall denote $h(\sigma)$, is obtained by taking the convolution of the pdf's of $Y_{n,t}$ and w_t, that, again dropping the subscripts, reads

$$h(\sigma) = \sum_{y>0} \int \delta(\sigma - wy) f(y) g(w)\, dw. \tag{6.10}$$

Introducing the auxiliary variable $z = wy$ and noting that $dz = ydw$ the last expression gives

$$h(\sigma) = \sum_{y>0} \int \frac{1}{y} \delta(\sigma - z) f(y) g\left(\frac{z}{y}\right) dz$$

$$= \sum_{y>0} \frac{f(y)}{y} g\left(\frac{\sigma}{y}\right). \tag{6.11}$$

Thus the pdf of the volatility σ_t is a weighted sum of the pdf of the market volatility w_t rescaled by the values that the process $Y_{n,t}$ can take, where weights are the probabilities of each of the discrete values y.

6.2.3 Moments

We can compute closed-form expressions for the moments of the distribution $h(\sigma)$. Let us start with the mean

$$\langle \sigma \rangle = \langle Y_{n,t} w_t \rangle \tag{6.12}$$

which, because of independence of $Y_{n,t}$ from w_t, can be factorized as

$$\begin{aligned} \langle \sigma \rangle &= \langle \prod_{m \in \gamma_i} e^{K_m} \rangle \langle w_t \rangle \\ &= \prod_{m \in \gamma_i} \langle e^{K_m} \rangle \langle w_t \rangle \\ &= \prod_{m \in \gamma_i} \zeta_1(p_m)\, e^{s^2/2}, \end{aligned} \tag{6.13}$$

since $\langle w_t \rangle = e^{s^2/2}$ and we define $\zeta_1(p_m) = \langle e^{K_m} \rangle = p_m(e-1)+1$. The moment of order n is defined as

$$m_n = (-i)^n \, \frac{d^n}{d\omega^n} \hat{h}(\omega) \Big|_{\omega=0}, \tag{6.14}$$

where $\hat{h}(\omega)$ is the characteristic function of $h(\sigma)$. This can be computed as

$$\begin{aligned}
\hat{h}(\omega) &= \int_{-\infty}^{+\infty} e^{i\omega\sigma} h(\sigma) d\sigma \\
&= \int_{-\infty}^{+\infty} e^{i\omega\sigma} \sum_{y>0} \frac{f(y)}{y} g\left(\frac{\sigma}{y}\right) d\sigma \\
&= \int_{-\infty}^{+\infty} e^{i\omega\sigma} \sum_{y>0} \frac{f(y)}{y} \frac{y}{s\sigma\sqrt{2\pi}} e^{-\frac{1}{2s^2}\ln^2(\sigma/y)} d\sigma \\
&= \sum_{y>0} \frac{f(y)}{s\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{i\omega\sigma} \, \frac{1}{\sigma} \, e^{-\frac{1}{2s^2}\ln^2(\sigma/y)} d\sigma.
\end{aligned} \tag{6.15}$$

After performing the change of variable $\tilde{\sigma} = \sigma/y$ with $d\tilde{\sigma} = d\sigma/y$ equation (6.15) becomes

$$\begin{aligned}
\hat{h}(\omega) &= \sum_{y>0} f(y) \frac{1}{s\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{i\omega y\tilde{\sigma}} \, \frac{1}{\tilde{\sigma}} \, e^{-\frac{1}{2s^2}\ln^2\tilde{\sigma}} d\tilde{\sigma}. \\
&= \sum_{y>0} f(y) \, \hat{g}(\omega y),
\end{aligned} \tag{6.16}$$

where $\hat{g}(\omega)$ is the characteristic function of the density $g(x)$,

$$\hat{g}(\omega) = \int_{-\infty}^{+\infty} e^{i\omega x} g(x) \, dx. \tag{6.17}$$

A representation for $\hat{g}(\omega)$ is know to be the following [48]

$$\hat{g}(\omega) = \sum_{n=0}^{\infty} \frac{(i\omega)^n}{n!} e^{n^2 s^2/2} \tag{6.18}$$

which allows to write

$$\hat{h}(\omega) = \sum_{y>0} f(y) \sum_{n=0}^{\infty} \frac{(i\omega\, y)^n}{n!} e^{n^2 s^2/2}. \tag{6.19}$$

From the last expression the order n moment m_n is obtained straightforwardly through Equation (6.14) as

$$m_n = \sum_{y>0} f(y)\; y^n e^{n^2 s^2/2}. \tag{6.20}$$

The variance $\mathrm{Var}(\sigma)$, in particular, can be computed using Equation (6.20) as

$$\begin{aligned} \mathrm{Var}(\sigma) =& m_2 - \langle\sigma\rangle^2 = \sum_{y>0} f(y) y^2 e^{2s^2} - \prod_{m\in\gamma_i} \zeta_1^2(p_m) e^{s^2} \\ =& e^{s^2} \left(\sum_{y>0} f(y) y^2 e^{s^2} - \prod_{m\in\gamma_i} \zeta_1^2(p_m) \right). \end{aligned} \tag{6.21}$$

6.2.4 Correlation Structure

The correlation matrix of the multivariate vector $\boldsymbol{r}_t$ crucially depends on the hierarchical factor $Y_{n,t}$. In this section we derive the exact expression for its entries.

Proposition 2. The correlation coefficient between two process $r_{i,t}$ and $r_{j,t}$ has the following form

$$\rho_{ij} = \mathrm{Corr}(r_i r_j) = \mathrm{Corr}(\epsilon_i \epsilon_j) \mathcal{F}_{ij}(\mathbf{p}; \Gamma_{i,t}, \Gamma_{j,t}) \tag{6.22}$$

where

$$\mathcal{F}_{ij}(\mathbf{p}; \Gamma_{i,t}, \Gamma_{j,t}) = \frac{\displaystyle\prod_{l:a_l\in\Gamma_{i,t}\setminus\Gamma_{j,t}} \zeta_1(p_l) \prod_{h:a_h\in\Gamma_{j,t}\setminus\Gamma_{i,t}} \zeta_1(p_h)}{\left(\displaystyle\prod_{l:a_l\in\Gamma_{i,t}\setminus\Gamma_{j,t}} \zeta_2(p_l) \prod_{h:a_h\in\Gamma_{j,t}\setminus\Gamma_{i,t}} \zeta_2(p_h) \right)^{1/2}}. \tag{6.23}$$

Proof. The correlation coefficient between two returns r_i and r_j is given by

$$\rho_{ij} = \frac{\mathrm{Cov}(r_i r_j)}{\sqrt{\mathrm{Var}(r_i)\mathrm{Var}(r_j)}}. \tag{6.24}$$

We drop the subscript t for the sake of readability and consider the two random variables r_i and r_j at arbitrary time t. Let us consider two return processes r_i and r_j having arbitrary hierarchical trees:

$$\Gamma_i = \{a_1, \ldots, a_{n_i}\} \tag{6.25}$$

$$\Gamma_j = \{b_1, \ldots, b_{n_j}\} \tag{6.26}$$

with corresponding risks

$$\Gamma_i \to K_{a_1}, \ldots, K_{a_{n_i}} \tag{6.27}$$

$$\Gamma_j \to K_{b_1}, \ldots, K_{b_{n_j}}. \tag{6.28}$$

Note that the Γ_i and Γ_j may or may not overlap. Denoting c the arbitrary node in the dendrogram, we consider also the sets

$$\begin{aligned}
\Gamma_i \cap \Gamma_j &= \{c : c \in \Gamma_i \wedge c \in \Gamma_j\} \\
\Gamma_i \setminus \Gamma_j &= \{c : c \in \Gamma_i \wedge c \notin \Gamma_j\} \\
\Gamma_j \setminus \Gamma_i &= \{c : c \in \Gamma_j \wedge c \notin \Gamma_i\}.
\end{aligned} \tag{6.29}$$

The covariance between r_i and r_j is

$$\mathrm{Cov}(r_i r_j) = \langle r_i r_j \rangle = \langle \epsilon_i \epsilon_j \rangle \langle x_t^2 \rangle \langle Y_{n_i} Y_{n_j} \rangle \tag{6.30}$$

$$= \langle \epsilon_i \epsilon_j \rangle \langle x_t^2 \rangle \prod_{x : a_x \in \Gamma_i \cap \Gamma_j} \langle e^{2K_x} \rangle \prod_{l : a_l \in \Gamma_i \setminus \Gamma_j} \langle e^{K_l} \rangle \prod_{h : a_h \in \Gamma_j \setminus \Gamma_i} \langle e^{K_h} \rangle. \tag{6.31}$$

Since $\langle K_m \rangle = p_m$, we have $\langle e^{2K_m} \rangle = p_m(e^2 - 1) + 1 = \zeta_2(p_m)$ and $\langle e^{K_m} \rangle = p_m(e - 1) + 1 = \zeta_1(p_m)$. One also has

$$\mathrm{Var}(r_i) = \langle \epsilon_i^2 \rangle \prod_{a \in \Gamma_i} \langle e^{2K_a} \rangle = \langle \epsilon_i^2 \rangle \prod_{m=1}^{n_i} \langle e^{2K_m} \rangle \tag{6.32}$$

$$\mathrm{Var}(r_j) = \langle \epsilon_i^2 \rangle \prod_{b \in \Gamma_j} \langle e^{2K_b} \rangle = \langle \epsilon_i^2 \rangle \prod_{m=1}^{n_j} \langle e^{2K_m} \rangle. \tag{6.33}$$

The correlation coefficient therefore is

$$\rho_{ij} = \mathrm{Corr}(\epsilon_i \epsilon_j) \frac{\prod\limits_{x:a_x \in \Gamma_i \cap \Gamma_j} \zeta_2(p_x) \prod\limits_{l:a_l \in \Gamma_i \setminus \Gamma_j} \zeta_1(p_l) \prod\limits_{h:a_h \in \Gamma_j \setminus \Gamma_i} \zeta_1(p_h)}{\left(\prod\limits_{m=1}^{n_i} \zeta_2(p_m) \prod\limits_{q=1}^{n_j} \zeta_2(p_m) \right)^{1/2}}. \tag{6.34}$$

The last expression can be further simplified by noting that

$$\prod_{m=1}^{n_i} \zeta_2(p_m) = \prod_{z:a_x \in \Gamma_i \cap \Gamma_j} \zeta_2(p_x) \times \prod_{l:a_l \in \Gamma_i \setminus \Gamma_j} \zeta_2(p_l) \tag{6.35}$$

and

$$\prod_{m=1}^{n_j} \zeta_2(p_m) = \prod_{z:a_x \in \Gamma_i \cap \Gamma_j} \zeta_2(p_x) \times \prod_{h:a_h \in \Gamma_j \setminus \Gamma_i} \zeta_2(p_h). \tag{6.36}$$

The correlation coefficient ρ_{ij} then reads

$$\rho_{ij} = \mathrm{Corr}(\epsilon_i \epsilon_j) \frac{\prod\limits_{x:a_x \in \Gamma_i \cap \Gamma_j} \zeta_2(p_x) \prod\limits_{l:a_l \in \Gamma_i \setminus \Gamma_j} \zeta_1(p_l) \prod\limits_{h:a_h \in \Gamma_j \setminus \Gamma_i} \zeta_1(p_h)}{\prod\limits_{z:a_x \in \Gamma_i \cap \Gamma_j} \zeta_2(p_x) \left(\prod\limits_{h:a_h \in \Gamma_j \setminus \Gamma_i} \zeta_2(p_h) \prod\limits_{l:a_l \in \Gamma_i \setminus \Gamma_j} \zeta_2(p_l) \right)^{1/2}} \tag{6.37}$$

which yields

$$\rho_{ij} = \mathrm{Corr}(\epsilon_i \epsilon_j) \frac{\prod\limits_{l:a_l \in \Gamma_i \setminus \Gamma_j} \zeta_1(p_l) \prod\limits_{h:a_h \in \Gamma_j \setminus \Gamma_i} \zeta_1(p_h)}{\left(\prod\limits_{h:a_h \in \Gamma_j \setminus \Gamma_i} \zeta_2(p_h) \prod\limits_{l:a_l \in \Gamma_i \setminus \Gamma_j} \zeta_2(p_l) \right)^{1/2}}. \tag{6.38}$$

Defining

$$\mathcal{F}_{ij}(\mathbf{p}; \Gamma_i, \Gamma_j) = \frac{\prod\limits_{l:a_l \in \Gamma_i \setminus \Gamma_j} \zeta_1(p_l) \prod\limits_{h:a_h \in \Gamma_j \setminus \Gamma_i} \zeta_1(p_h)}{\left(\prod\limits_{l:a_l \in \Gamma_i \setminus \Gamma_j} \zeta_2(p_l) \prod\limits_{h:a_h \in \Gamma_j \setminus \Gamma_i} \zeta_2(p_h) \right)^{1/2}} \tag{6.39}$$

we can therefore write

$$\rho_{ij} = \mathrm{Corr}(\epsilon_i \epsilon_j) \mathcal{F}_{ij}(\mathbf{p}; \Gamma_i, \Gamma_j). \tag{6.40}$$

□

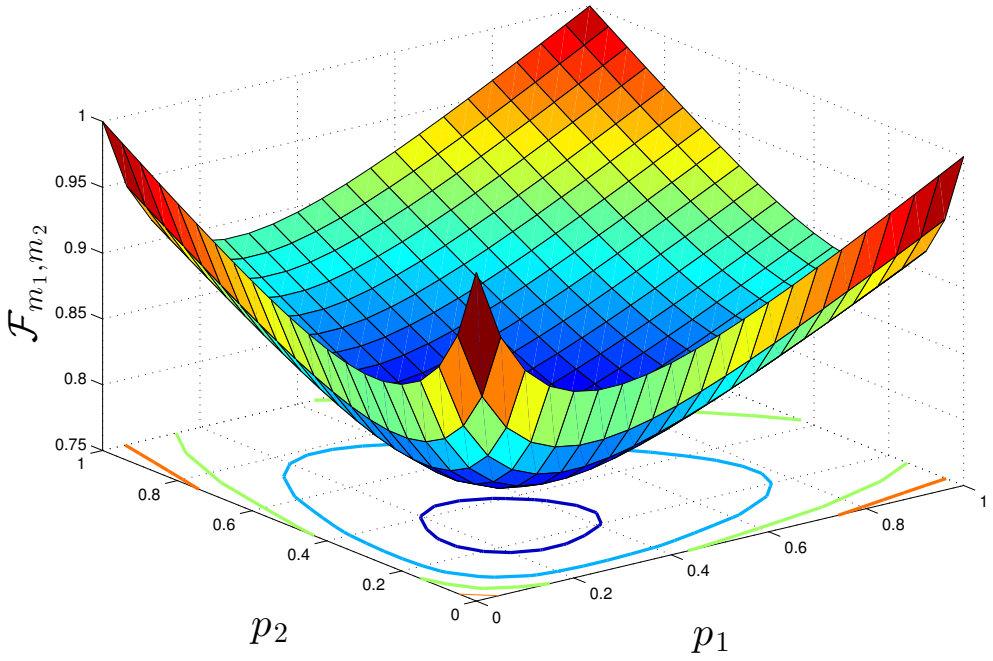

FIGURE 6.5: We plot the surface $\mathcal{F}_{m_1,m_2}$ in Equation (6.41) as a function of the probabilities p_{m_1} and p_{m_2} associated with the arbitrary nodes m_1 and m_2.

Some important remarks:

- The total correlation is factorized into two separate contributions, the first one associated with the multivariate Gaussian vector $\boldsymbol{\epsilon}_t$ and the second one only dependent on the probabilities of the nodes, that is on the hierarchy.
- The hierarchy factor $\mathcal{F}_{ij}(\mathbf{p};\Gamma_i,\Gamma_j)$ is always smaller than one and thus acts as a perturbation of $\mathrm{Corr}(\epsilon_i\epsilon_j)$, damping it according to the values assumed by the parameters p's.
- The two limit cases $p_m = 0$, $\forall m = 1,\ldots,N-1$ and $p_m = 1$, $\forall m = 1,\ldots,N-1$ both correspond to the hierarchy factor being 1. In both cases the effect of the hierarchy is null and one has $\rho_{i,j} = \mathrm{Corr}(\epsilon_i,\epsilon_j)$ (see Appendix H).

Overall one can conclude that including a hierarchical structure in the volatility modelling introduces a perturbation to the standard multivariate model correlation matrix, which is nonetheless recovered when the probabilities of the internal

risks are all null or all unity. These two cases correspond indeed to have the hierarchical structure disappearing, since there would not be any difference between different trees. The limit $\mathcal{F}_{ij}(\mathbf{p};\Gamma_i,\Gamma_j)=1$ corresponds in fact to the case where the market is dominated by one single risk and the multi-branched hierarchical structure disappears.

In order to visualise the behaviour of the hierarchical factor $\mathcal{F}_{ij}(\mathbf{p};\Gamma_i,\Gamma_j)$ with the probabilities of the internal nodes let us consider the two dimensional case

$$\mathcal{F}_{m_1,m_2}=\frac{\zeta_1(p_{m_1})\zeta_1(p_{m_2})}{\left(\zeta_2(p_{m_1})\zeta_2(p_{m_2})\right)^{1/2}} \tag{6.41}$$

which is plotted in Figure 6.5. Intuitively one can guess how adding more risks will result in the hyper-surface $\mathcal{F}_{ij}(\mathbf{p};\Gamma_i,\Gamma_j)$ attaining smaller values in $[0,1]$. One can also appreciate that the hierarchical factor $\mathcal{F}_{ij}(\mathbf{p};\Gamma_i,\Gamma_j)$ is 1 at the four corners and <1 for all other combinations of the probabilities, resulting in the overall correlation coefficient being shrunk.

We show in the left plot in Figure 6.6 the distribution of measured correlations on a multivariate synthetic DHM with probabilities $p_m,\ m=1,\dots N-1$ uniformly distributed in the range $[0.1,0.4]$ compared to a multivariate log-normal model, which is recovered from the DHM when all p_m's are one. One can see how the distribution in the case of $\mathcal{F}_{ij}(\mathbf{p};\Gamma_i,\Gamma_j)\neq 1$ is more skewed to the left. Increasing the values of the probabilities progressively shifts the median to the right, as shown in the right plot of Figure 6.6, where the distribution of observed correlation coefficients is plotted on different realisations of DHM with probabilities taking values in different ranges. We can see how increasing the range of probability values shifts the bulk of the distribution to the right, which eventually matches the one of a log-normal multivariate model, where the hierarchical factors is turned off.

6.2.5 Autocorrelation Function

The hierarchical structure affects likewise the autocorrelation structure of the single processes $r_{i,t}$, for $i=1,\dots,N$. Intuitively, if the process w_t is long-range

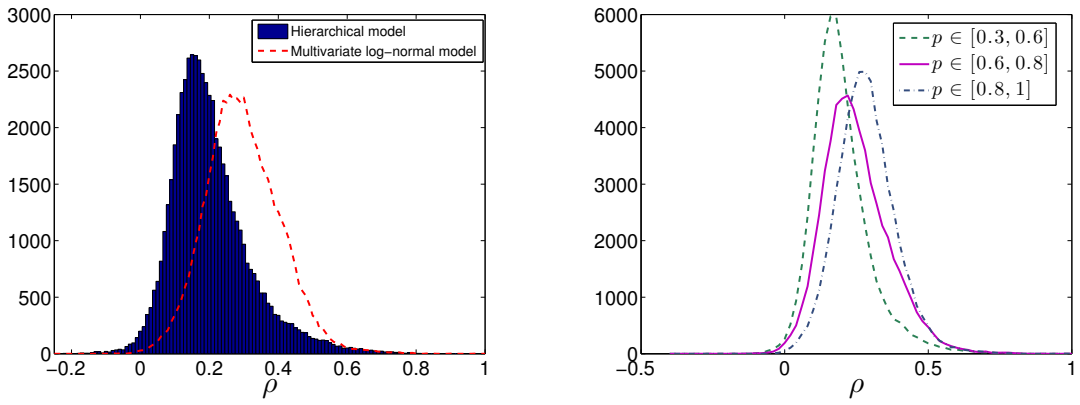

FIGURE 6.6: (Left) Distribution of the observed correlation coefficients on DHM model simulated with probability randomly drawn in $[0.1, 0.4]$ (blue bars) compared to the case where all p's are 1 (red dashed line). (Right) Distributions of the observed correlation coefficients on simulated DHM multivariate time series with probabilities of the internal risks in different ranges, reported in the legend.

correlated, the convolution with the (uncorrelated) process $Y_{n,t}$ will result in a reduction of the correlation function. We assume $Y_{n,t}$ to be uncorrelated in time, although further analysis in the time dependence of the market hierarchical structure is revealing interesting non-trivial patterns which suggest that even the clusters organisation is correlated in time [115]. We leave the case of time-correlated $Y_{n,t}$ for further investigations.

The volatility autocorrelation function is given by

$$C_\sigma(\ell) = \frac{\mathrm{Cov}(\sigma_t \sigma_{t+\ell})}{\sqrt{\mathrm{Var}(\sigma_t)\mathrm{Var}(\sigma_{t+\ell})}}, \tag{6.42}$$

where, as usual,

$$\mathrm{Cov}(\sigma_t \sigma_{t+\ell}) = \langle \sigma_t \sigma_{t+\ell} \rangle - \langle \sigma_t \rangle \langle \sigma_{t+\ell} \rangle. \tag{6.43}$$

In the evaluation of the mixed term $\langle \sigma_t \sigma_{t+\ell} \rangle$ one has to take into account the overlap of the hierarchy at time t with that at time $t + \ell$. Using the same notation already introduced in the previous section to evaluate the cross-correlation

coefficients, one has

$$\begin{aligned}
\langle \sigma_t \sigma_{t+\ell} \rangle &= \langle w_t w_{t+\ell} \rangle \langle \prod_{m=1}^{n_t} e^{K_{m,t}} \prod_{p=1}^{n_{t+\ell}} e^{K_{p,t+\ell}} \rangle \\
&= \langle w_t w_{t+\ell} \rangle \langle \exp\left(\sum_{m=1}^{n_t} K_{m,t} \right) \exp\left(\sum_{p=1}^{n_{t+\ell}} K_{p,t+\ell} \right) \rangle \\
&= \langle w_t w_{t+\ell} \rangle \langle \exp\left(2 \sum_{m: a_m \in \Gamma_t \cap \Gamma_{t+\ell}} K_m \right) \cdot \exp\left(\sum_{p: a_p \in \Gamma_t \setminus \Gamma_{t+\ell}} K_p \right) \cdot \\
&\quad \cdot \exp\left(\sum_{q: a_q \in \Gamma_{t+\ell} \setminus \Gamma_t} K_q \right) \rangle \\
&= \langle w_t w_{t+\ell} \rangle \prod_{m:\, a_m \in \Gamma_t \cap \Gamma_{t+\ell}} \langle e^{2K_m} \rangle \cdot \prod_{p:\, a_p \in \Gamma_t \setminus \Gamma_{t+\ell}} \langle e^{K_p} \rangle \cdot \prod_{q:\, a_q \in \Gamma_{t+\ell} \setminus \Gamma_t} \langle e^{K_q} \rangle \\
&= \langle w_t w_{t+\ell} \rangle \prod_{m:\, a_m \in \Gamma_t \cap \Gamma_{t+\ell}} \zeta_2(p_m) \cdot \prod_{p:\, a_p \in \Gamma_t \setminus \Gamma_{t+\ell}} \zeta_1(p_p) \cdot \prod_{q:\, a_q \in \Gamma_{t+\ell} \setminus \Gamma_t} \zeta_1(p_q).
\end{aligned} \tag{6.44}$$

It follows that, since w_t is stationary, the auto-covariance of σ_t is given by

$$\begin{aligned}
\mathrm{Cov}(\sigma_t \sigma_{t+\ell}) &= \langle w_t w_{t+\ell} \rangle \prod_{m:\, a_m \in \Gamma_t \cap \Gamma_{t+\ell}} \zeta_2(p_m) \cdot \prod_{p:\, a_p \in \Gamma_t \setminus \Gamma_{t+\ell}} \zeta_1(p_p) \cdot \prod_{q:\, a_q \in \Gamma_{t+\ell} \setminus \Gamma_t} \zeta_1(p_q) \\
&\quad - \langle w_t \rangle^2 \prod_{j=1}^{n_t} \zeta_1(p_j) \cdot \prod_{k=1}^{n_{t+\ell}} \zeta_1(p_k).
\end{aligned} \tag{6.45}$$

The variance of σ_t is instead given by

$$\mathrm{Var}(\sigma_t) = \langle \sigma_t^2 \rangle - \langle \sigma_t \rangle^2 = \langle w_t^2 \rangle \prod_{j=1}^{n_t} \zeta_2(p_j) - \langle w_t \rangle^2 \left(\prod_{j=1}^{n_t} \zeta_1(p_j) \right)^2. \tag{6.46}$$

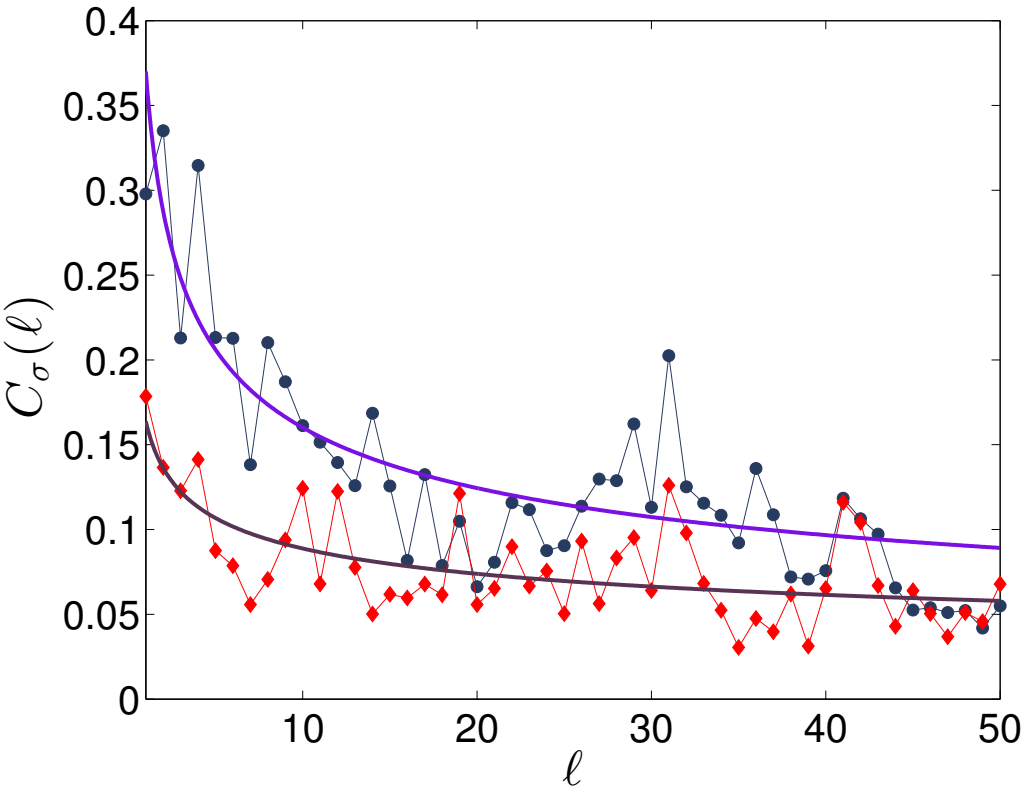

FIGURE 6.7: The autocorrelation function of a simulated time series r_t with ten risks and p uniformly drawn in $[0, 1]$ (red diamonds) is compared to that of a log-normal MRW process (blue circles). Note that the latter corresponds to a DHM process where $p_m = 0$ for $m = 1, \ldots, 10$. The thick lines are the two power-law fits with decay exponents $\beta = 0.26$ for DHM and $\beta = 0.37$ for MRW.

Combining Equations (6.45) and (6.46) we have

$$C_\sigma(\ell) = \frac{\langle w_t w_{t+\ell} \rangle \prod_{m:\, a_m \in \Gamma_t \cap \Gamma_{t+\ell}} \zeta_2(p_m) \cdot \prod_{p:\, a_p \in \Gamma_t \setminus \Gamma_{t+\ell}} \zeta_1(p_p) \cdot \prod_{q:\, a_q \in \Gamma_{t+\ell} \setminus \Gamma_t} \zeta_1(p_q)}{\sqrt{\mathrm{Var}(\sigma_t)\mathrm{Var}(\sigma_{t+\ell})}} - \frac{\langle w_t \rangle^2 \prod_{j=1}^{n_t} \zeta_1(p_j) \cdot \prod_{k=1}^{n_{t+\ell}} \zeta_1(p_k)}{\sqrt{\mathrm{Var}(\sigma_t)\mathrm{Var}(\sigma_{t+\ell})}}. \tag{6.47}$$

We can ascertain the effect of $Y_{n,t}$ on the factor w_t autocorrelation function by plotting the $C_\sigma(\ell)$ and $C_w(\ell)$ measured on simulated DHM time series. As shown in Figure 6.7, $Y_{n,t}$ does not spoil the long range nature of the volatility autocorrelation, although it modifies it, particularly for small lags. Values found for the power-law decay exponent β are nevertheless comparable to those observed on empirical data, that is in the interval $[0.1, 0.6]$.

Note that in the limit cases $p_m = 0, 1$ for $m = 1, \ldots, N-1$ one recovers the autocorrelation function $C_w(\ell)$.

So far we have discussed how the inclusion of the hierarchical term $Y_{n,t}$ in the volatility affects univariate properties (moments and autocorrelation function) as well as multivariate ones (correlation structure). We next look at the most relevant aspect in the context of this doctoral thesis, that is the effect on the multi-scaling properties of the time series.

6.3 Effect of the Hierarchy on the Scaling Properties

In this paragraph we study how the scaling properties of the volatility w_t get modified by the effect of the hierarchical factor $Y_{n,t}$. As we have seen before, $Y_{n,t}$ affects both the tails of the returns distribution and the autocorrelation function, the two main causes of multi-scaling behaviour.

We investigate the cumulative sum of the simulated return process, which we denote $X_{i,t} = \sum_{t'=1}^{t} r_{t',i}$. In order to simulate the MDHM time series, we first need to choose a particular correlation hierarchical structure: this is recovered via the DBHT clustering algorithm [22]. We first consider one constant hierarchical structure for the entire time period and we specify to a set of 25 stocks (see figure 6.8) taken from the set of 342 most capitalised firms traded in NYSE in the period 1995-2012[1]. Names of the stocks used are reported in Appendix I. Note that the clustering algorithm returns a hierarchy with $N - 1 = 24$ nodes. To each node, corresponding to a specific risk affecting all stocks sharing that node, we associate a random probability uniformly distributed in $[0, 1]$.

Choosing w_t as a log-normal process with power-law auto covariance function, one can recover the MRW [20] from our process $r_{t,i}$ by sending $p_m \to 0,\ m = 1, \ldots, 24$. This corresponds to $Y_{n,t} = 1$ in equation (6.3), i.e. the effect of the hierarchical

[1]The choice of this set of stocks is arbitrary and does not represent any loss of generality, but allows to explain the simulation in a simple fashion. Other choices have also been extensively investigated.

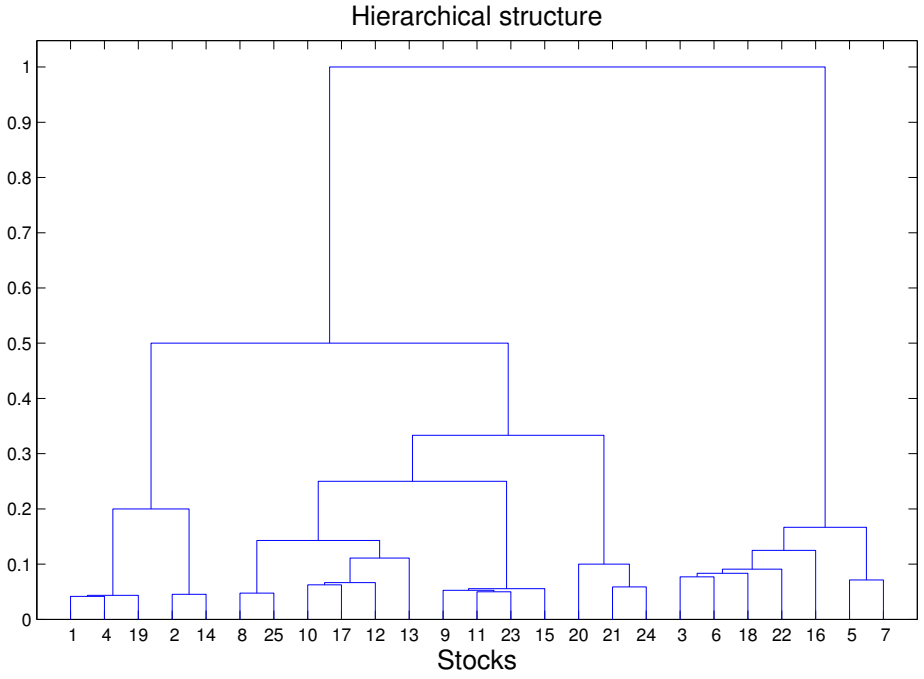

FIGURE 6.8: DBHT cross-correlation hierarchical structure for a set of 25 stocks, whose name and tickers are reported in Appendix I. The number of risks, corresponding to the nodes of the dendrogram, is $N - 1 = 24$.

structure is turned off. In this last case returns reduce, for each variable i, to $r_{t,i} = \epsilon_{i,t} w_t$ which, for the choice of w_t mentioned above, is the multivariate version of the multifractal process of Bacry *et al* [24] discussed in Section 5.4.1. The hierarchical structure is thus an addendum whose effect on the process is completely separable but can trigger, as we now see, interesting effects on the scaling properties of the aggregated process.

We first look at scaling of the moments of the increments of the process, namely (dropping the index i) [13]

$$M_q(t, \tau) = \langle |X_{t+\tau} - X_t|^q \rangle. \tag{6.48}$$

The scaling of the moments for $q \in [1, 3]$ is reported in Figures 6.9 and 6.10. We note a good scaling up to $q = 2$ which is thereupon progressively lost increasing q. This behaviour is common for a large class of asset prices [38, 39] (including stocks) and is also common to the MRW (see figure 6.11).

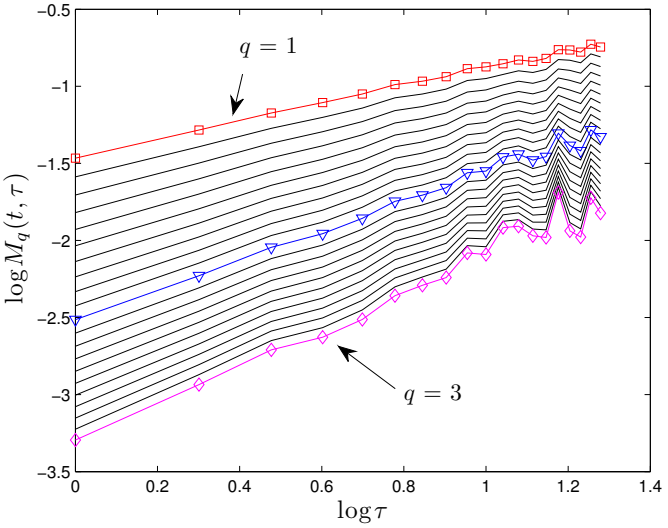

FIGURE 6.9: $M_q(t,\tau)$ as a function of τ in log-log scale for a realisation of the DHM. Each black curve corresponds to a different value of q ranging in [1,3]. The curves corresponding to $q = 1, 2, 3$ are highlighted as $(\square), (\triangledown)$ and $(\diamond)$ respectively.

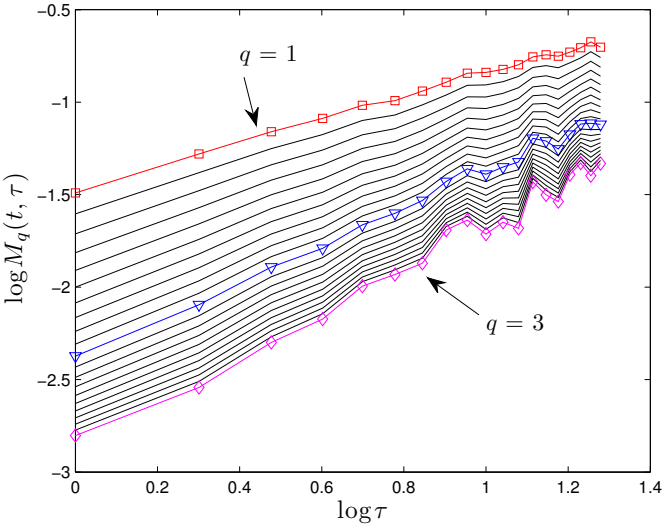

FIGURE 6.10: $M_q(t,\tau)$ as a function of τ in log-log scale for another realisation of the DHM. Each black curve corresponds to a different value of q ranging in [1,3]. The curves corresponding to $q = 1, 2, 3$ are highlighted as $(\square), (\triangledown)$ and $(\diamond)$ respectively.

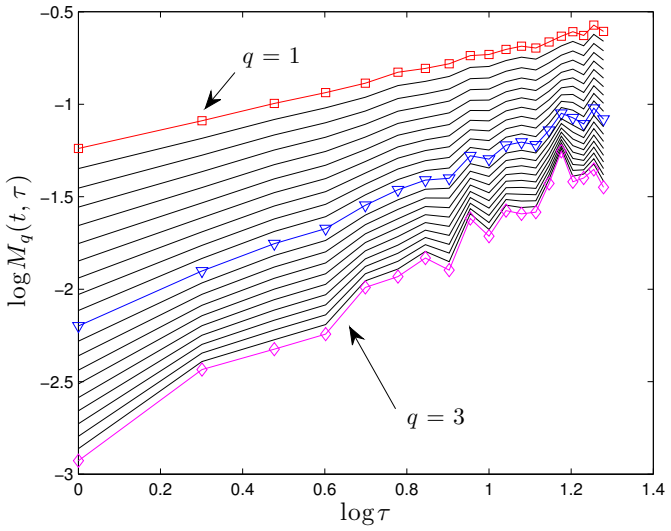

FIGURE 6.11: $M_q(t,\tau)$ as a function of τ in log-log scale for a realisation of the MRW. Each black curve corresponds to a different value of q ranging in [1,3]. The curves corresponding to $q = 1, 2, 3$ are highlighted as $(\square), (\triangledown)$ and $(\diamond)$ respectively.

The effect of $Y_{n,t}$ on the scaling properties of the process can be quantified comparing the scaling of the empirical moments (and the relevant Generalised Hurst Exponents) for the aggregate process X_t obtained by adding up MRW returns and that for the aggregate process obtained from the DHM returns. In Figure 6.12 we show two examples. Again, the processes have been simulated by considering random probabilities drawn uniformly in $[0, 1]$ for each node and a log-normal long-ranged correlated process for the volatility component w_t. Note that the process w_t itself is enough to have a concave scaling function ζ_q. The effect of the $Y_{n,t}$ on the returns dynamics is however clearly visible, insofar as the flexing of the scaling function is significantly enhanced. The hierarchical factor adds complexity to the price dynamics and this complexity is reflexed in the increased multifractal properties of the simulated processes.

6.3.1 Time-varying Multifractality

As a consequence of their intrinsic non-stationarity, multivariate financial data also show time varying cross-correlation hierarchical structure [115]. Especially over turbulent market periods, as we have discussed in Chapter 4, one recovers

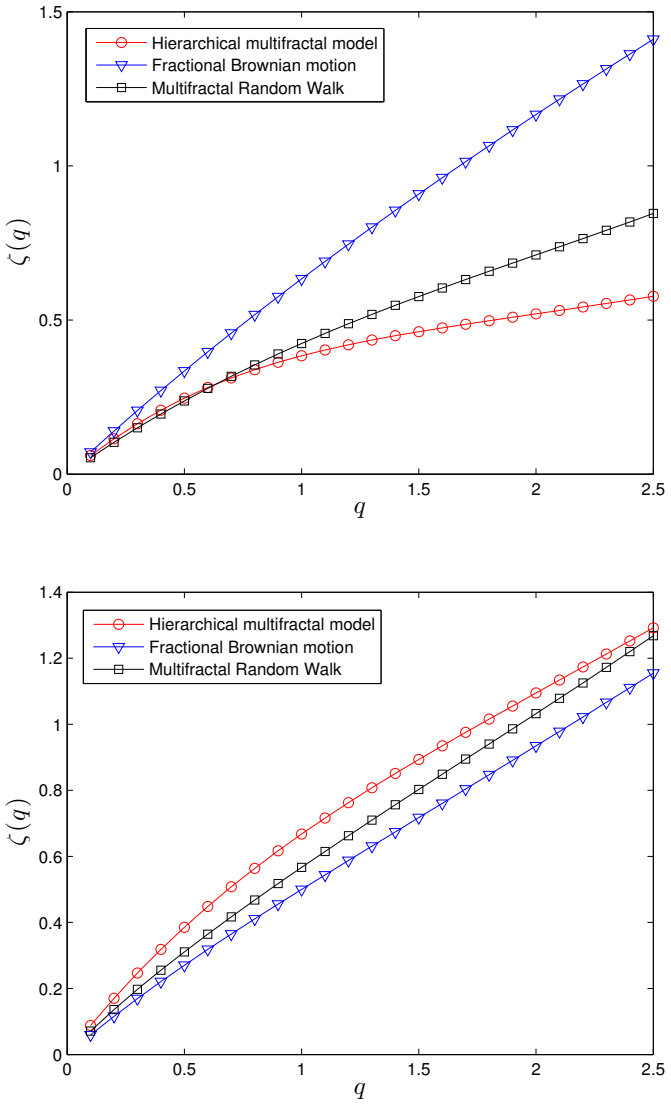

FIGURE 6.12: An example of the scaling function ζ_q of a realisation of the DHM (red circles) and of a realisation of a MRW process (black squares). We also plot the scaling of a a Fractional Brownian motion (blue triangles) with Hurst parameter $H = 0.6$, whose scaling is linear with q.

very different organisations when studying the hierarchical properties on moving time windows. Since the MDHM setting accounts for time-varying hierarchy of correlations, it is natural to associate the time-varying multifractal properties of the stock market reported in Chapter 3 with a time-varying hierarchical structure. To this end we have performed DBHT clustering dynamically over non-overlapping time windows and looked at the scaling properties of simulated MDHM time series with a time varying hierarchical structure. As an illustrative case, being also the

most straightforward to consider, we have split the 25 NYSE time series into two tranches (regimes) of equal length: the first tranche spans the period 1/1/1995 through 31/12/2004, the second one 31/12/2004 through 22/10/2012. In this way, both tranches contain 2013 price observations. We label the two regimes T_1 and T_2 and the corresponding hierarchies $\mathcal{H}_{T_1}$ and $\mathcal{H}_{T_2}$. We report the dendrograms corresponding to the two hierarchies $\mathcal{H}_{T_1}$ and $\mathcal{H}_{T_2}$ in Appendix J. In regime T_1 the hierarchical order n_{T_1} ranges in $[3, 10]$ while in regime T_2 it ranges in $n_{T_2} \in [2, 8]$. To each stock $i = 1, \ldots, 25$ is associated a hierarchical tree $\Gamma_i^{T_k}$, for $k = 1, 2$, including all the nodes above the stock. We plot in Figure 6.13 the simulated returns and volatility for the stock labelled 11 in Figure J.1, as an example of the effect of the changing hierarchy. In the tranche T_1 this particular stock shares one common risk with stock labelled 14, then one common risk with the pair of stocks $6, 15$ and so on. In the tranche T_2 it shares one common risk with stock 25, one common risk with the pair $3, 9$ and so on. Overall the topology of the dendrogram changes quite dramatically in the second tranche and this result in a rather different spreading of the risks. For the stock labelled 11 the number of risks increase from 5 to 8. One can clearly see the effect of the larger number of risks which cause fluctuations to become much larger.

The multifractality measured dynamically on simulated time series with the two-regimes hierarchy extracted from the data shows fluctuations in time. These fluctuations in time can be compared with those expected to be observed on DHM with hierarchy $\mathcal{H}_1^{11}$ or on DHM with hierarchy $\mathcal{H}_2^{11}$. We show that fluctuations observed in the two regimes of the time series are compatible with the confidence interval expected for the DHM model with the corresponding hierarchy. We simulate many realisations of the DHM with two different hierarchical structure: the first one having hierarchy $\mathcal{H}_1^{11}$ and the second one having hierarchy $\mathcal{H}_2^{11}$. Computing $\Delta H^w(1, 2)$ each time, we obtain a two sample populations of observed $\Delta H^w(1, 2)$'s. The $\{2.5\%, 50\%, 97.5\%\}$-quantiles of the $\Delta H^w(1, 2)$ distribution observed on $\mathcal{H}_1^{11}$ are reported in the left plot of figure 6.14, whereas those observed on $\mathcal{H}_2^{11}$ are reported in the right plot. The figures show that, although the $\Delta H^w(1, 2)$ fluctuations observed in the second tranche of the time series are exceeding the

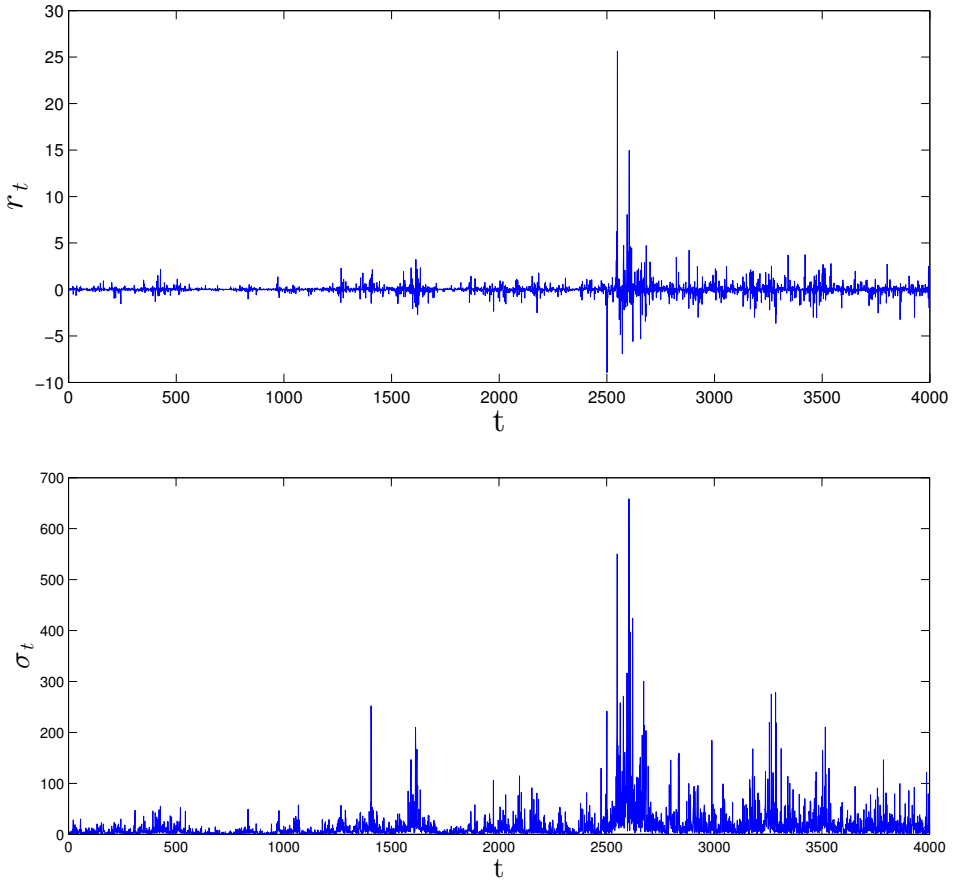

FIGURE 6.13: We plot returns (top) and volatility (bottom) for the simulated process whose hierarchical structure is $\Gamma_{11} = \Gamma_{11}^{T_1}$ for $t \in [1, 2013]$ and $\Gamma_{11} = \Gamma_{11}^{T_2}$ for $t \in [2014, 4026]$. The common volatility process x_t has been simulated as discussed in the Methods section of the main text with parameters $\lambda = 0.2$ and $T = 800$.

quantiles computed from time series generated with the hierarchy $\mathcal{H}_1^{11}$, they stay within the quantiles computed from time series generated with the hierarchy $\mathcal{H}_2^{11}$. In other words the fluctuations observed on multifractality in the second regime of hierarchy are due only to the switch between different hierarchies and one can account for them by considering the larger confidence intervals envisaged by the model with hierarchical structure given by $\mathcal{H}_2^{11}$.

We have performed an extensive analysis studying the measured multifractality and scaling exponents on many synthetic time series, considering separately those whose hierarchical order increases from those whose hierarchical order decreases. We have simulated 1000 realisations of a 25 stocks 2-regimes multivariate DHM of length 4000 with hierarchical structures given by $\mathcal{H}_{T_1}$ and $\mathcal{H}_{T_2}$ on the two tranches

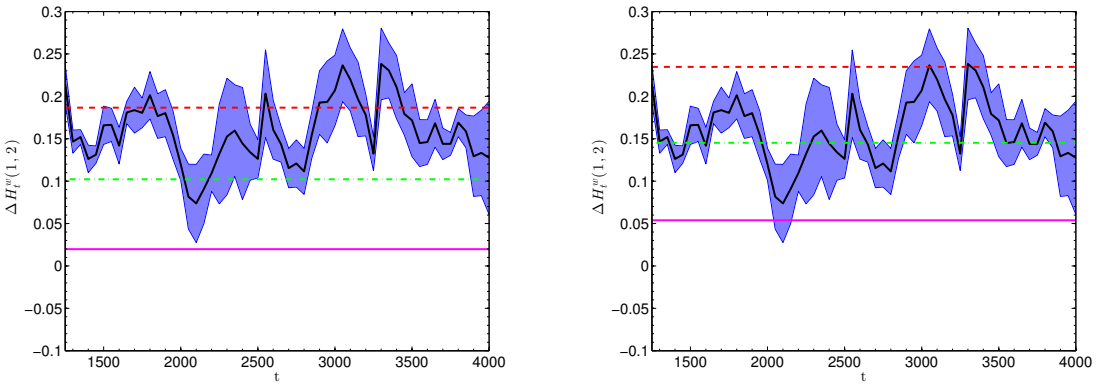

FIGURE 6.14: $\Delta H^w(1,2)$ dynamically computed on a realisation of the DHM with 2-regimes hierarchical structure given by $\mathcal{H}_1^{11}$ and $\mathcal{H}_2^{11}$. We also show in the plot on the left the $\{2.5\%, 50\%, 97.5\%\}$-quantiles obtained from a Monte Carlo simulation of 1000 realisations of 2000 time steps DHM with fixed hierarchical structure $\mathcal{H}_1^{11}$. On the right, the quantiles shown are those computed from a Monte Carlo simulation of 1000 realisations of 2000 time steps DHM with fixed hierarchical structure $\mathcal{H}_2^{11}$. In both figures the magenta line corresponds to the 2.5% quantiles, the green dot-dashed line to the 50% quantile and the red dashed line to the 97.5% quantile.

T_1 ($t \in [1, 2000]$) and T_2 ($t \in [2001, 4000]$) respectively. Each time the scaling exponents and multifractality indicator $\Delta H(1,2)$ have been computed, keeping track of whether $n_{T_1} > n_{T_2}$ or $n_{T_1} < n_{T_2}$ for each stock i. With this analysis one can test whether there is a significant correlation between the hierarchical order and the scaling properties of the synthetic time series. For the hierarchies in Figure J.1 we found 11 stocks whose order increases, 10 stocks whose order decreases and 4 stocks whose order does not change. This analysis has been performed for different values of the probabilities of the risks considering all p's to be uniformly distributed in $[0.4, 0.6]$, $[0.6, 0.8]$ and $[0.8, 1]$ respectively. The three sets of results are shown in Tables 6.2, 6.3 and 6.4. For each realisation of MDHM the scaling exponents have been averaged over the 25 stocks for the three different cases separately. Then, the values reported in Tables 6.2, 6.3 and 6.4 are the averages over the 1000 realisations. This allows to have a robust indication of the correlation between scaling properties and hierarchical order in time. From these tables two general facts stand out:

- the observed multifractality is correlated with the order of the hierarchy. In

	Increasing order		Decreasing order	
	T_1	T_2	T_1	T_2
H(1)	0.58	0.62	0.61	0.60
H(2)	0.48	0.51	0.48	0.51
$\Delta H(1,2)$	0.10	0.12	0.13	0.09

TABLE 6.2: Scaling exponents $H(1), H(2)$ and $\Delta H(1,2)$ obtained as averages with error over 1000 realisations of DHM with the two regimes hierarchical structures: $\mathcal{H}_{T_1}$ on the tranche T_1 and $\mathcal{H}_{T_2}$ on the tranche T_2. The probabilities of the risks have been chosen to be uniformly distributed in $[0.4, 0.6]$. Standard errors are of order 10^{-3} and not reported. All mean values in T_1 have been statistically checked to be significantly different from those in T_2 via a t-test, returning p-value always negligible at a threshold of 5% confidence level.

	Increasing order		Decreasing order	
	T_1	T_2	T_1	T_2
H(1)	0.55	0.60	0.58	0.57
H(2)	0.48	0.51	0.48	0.51
$\Delta H(1,2)$	0.07	0.09	0.10	0.06

TABLE 6.3: Scaling exponents $H(1), H(2)$ and $\Delta H(1,2)$ obtained as averages with error over 1000 realisations of DHM with the two regimes hierarchical structures: $\mathcal{H}_{T_1}$ on the tranche T_1 and $\mathcal{H}_{T_2}$ on the tranche T_2. The probabilities of the risks have been chosen to be uniformly distributed in $[0.6, 0.8]$. Standard errors are of order 10^{-3} and not reported.

	Increasing order		Decreasing order	
	T_1	T_2	T_1	T_2
H(1)	0.53	0.56	0.54	0.55
H(2)	0.48	0.51	0.48	0.51
$\Delta H(1,2)$	0.05	0.05	0.06	0.04

TABLE 6.4: Scaling exponents $H(1), H(2)$ and $\Delta H(1,2)$ obtained as averages with error over 1000 realisations of DHM with the two regimes hierarchical structures: $\mathcal{H}_{T_1}$ on the tranche T_1 and $\mathcal{H}_{T_2}$ on the tranche T_2. The probabilities of the risks have been chosen to be uniformly distributed in $[0.8, 1]$. Standard errors are of order 10^{-3} and not reported.

all cases indeed we observed higher multifractality on the tranches where the hierarchical order was larger than in those where it was smaller. As one can read from the three tables, simulated time series whose order increase switching from the first to the second hierarchy show increasing multifractality, whereas series whose order decreases show decreasing multifractality.

- Multifractality is largest for parameters p's in $[0.4, 0.6]$ and decreases significantly in the other two sets of simulations. The increase or decrease of multifractality and scaling exponents $H(1)$ and $H(2)$ have been statistically validated through t-test at two-sided 5% confidence level: the p-values are always extremely small, confirming a true differentiation between the scaling properties in T_1 and T_2.

A more complete assessment of the actual time-varying multifractal properties of the MDHM can be attained by looking at the quantiles of the $\Delta H(1,2)$ distribution over the two regimes. As we showed in Figure 6.14 for one example, quantiles shift upwards when large fluctuations kick in. This effect is observed robustly and coherently on the entire set of 25 simulated processes, using the two-regime hierarchical structure of Figure J.1. We plot in Figure 6.15 the q-quantiles $Q_q^{T_1}$ and $Q_q^{T_2}$ for $q = \{2.5\%, 50\%, 97.5\%\}$ of the distribution of $\Delta H(1,2)$ measured on the simulated DHM time series with hierarchies $\mathcal{H}_{T_1}$ in the first tranche and $\mathcal{H}_{T_2}$ in the second. Probabilities of the risks have been uniformly drawn in $[0.4, 0.6]$. Comparing the hierarchical orders of the stocks in both tranches, we plot in different colours the quantiles on the two different time windows, both for stocks whose hierarchical order increases (top) and those whose hierarchical order decreases (bottom). We observe a systematic shift of the distribution of $\Delta H(1,2)$ towards larger values for those stocks whose hierarchical order increases and a shift towards smaller values for those whose hierarchical order decreases. This confirms further that the underlying structure of risks can be a good candidate to explain the empirical observations presented in Chapter 4 as well as empirical observations

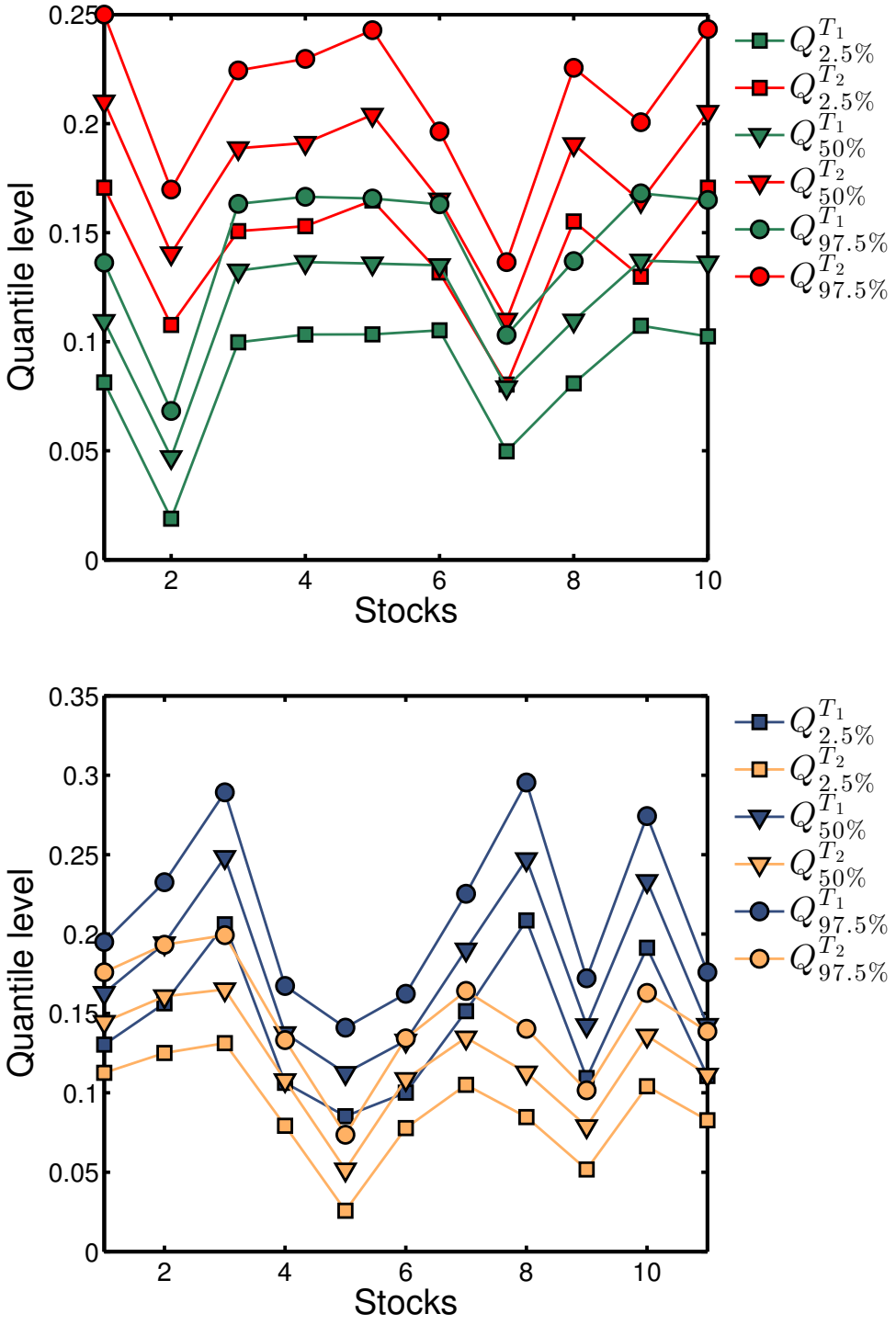

FIGURE 6.15: (Top) The plot shows the q-quantiles $Q_q^{T_1}$ and $Q_q^{T_2}$ for $q = \{2.5\%, 50\%, 97.5\%\}$ for the 10 DHM simulated time series whose hierarchical order $n_{T_1} < n_{T_2}$. Red lines correspond to p-quantiles in T_1 while black lines to p-quantiles in T_2. (Bottom) The q-quantiles $Q_q^{T_1}$ and $Q_q^{T_2}$ for $q = \{2.5\%, 50\%, 97.5\%\}$ for the 11 DHM simulated time series whose hierarchical order $n_{T_1} > n_{T_2}$. Blue lines correspond to q-quantiles in T_1 while orange lines to q-quantiles in T_2. The probabilities of the risks are initialised to values ranging in $[0.4, 0.6]$.

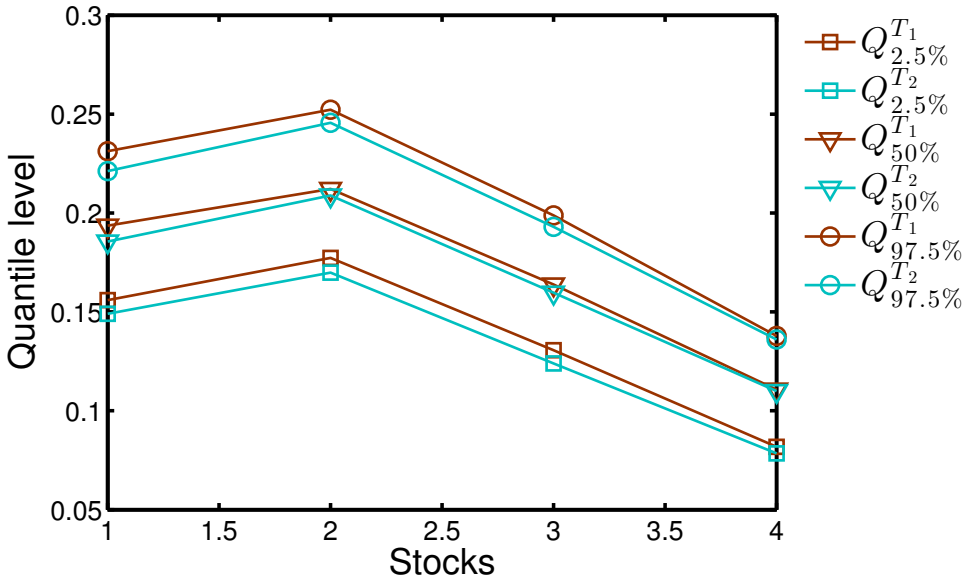

FIGURE 6.16: The plot shows the q-quantiles $Q_q^{T_1}$ and $Q_q^{T_2}$ for $q = \{2.5\%, 50\%, 97.5\%\}$ for the 4 DHM simulated time series whose hierarchical order $n_{T_1} = n_{T_2}$. Red lines correspond to p-quantiles in T_1 while light blue lines to p-quantiles in T_2.

of time-varying multifractality reported in Chapter 3. For the 4 stocks whose hierarchical order $n_{T_1} = n_{T_2}$ we show in Figure 6.16 the distribution of $\Delta H(1,2)$ obtained on the two different tranches. Compared to the plots in Figure 6.15, we can see how the difference between the values obtained in the two windows is much smaller for the four stocks with $n_{T_1} = n_{T_2}$. The quantiles are also reported in Tables 6.5, 6.6 and 6.7 for the three cases $n_{T_1} < n_{T_2}$, $n_{T_1} > n_{T_2}$ and $n_{T_1} = n_{T_2}$ respectively. Results of simulations with different probabilities are reported in Appendix K.

6.3.2 Multifractality depends on the hierarchy

As we have seen in Chapter 4, financial data show non trivial dependence between multifractality and cross-correlation hierarchical order. The MDHM has therefore been constructed with the aim to correlate the degree of multifractality with the hierarchical order of the process $r_{t,i}$. The collective action of the many risks affecting the stocks elicits a thickening of the tails which results in the process more affected being more multifractal. This establishes a direct relationship between

$n_{T_1} < n_{T_2}$	
T_1	T_2
$[Q^{T_1}_{2.5\%}, Q^{T_1}_{50\%}, Q^{T_1}_{97.5\%}]$	$[Q^{T_2}_{2.5\%}, Q^{T_2}_{50\%}, Q^{T_2}_{97.5\%}]$
$[0.0814, 0.1095, 0.1362]$	$[0.1706, 0.2101, 0.2499]$
$[0.0188, 0.0470, 0.0683]$	$[0.1077, 0.1405, 0.1697]$
$[0.0998, 0.1326, 0.1632]$	$[0.1507, 0.1887, 0.2244]$
$[0.1033, 0.1365, 0.1664]$	$[0.1530, 0.1911, 0.2297]$
$[0.1033, 0.1358, 0.1657]$	$[0.1647, 0.2040, 0.2428]$
$[0.1052, 0.1351, 0.1630]$	$[0.1317, 0.1651, 0.1964]$
$[0.0498, 0.0790, 0.1031]$	$[0.0804, 0.1101, 0.1365]$
$[0.0809, 0.1097, 0.1370]$	$[0.1551, 0.1905, 0.2256]$
$[0.1073, 0.1371, 0.1681]$	$[0.1298, 0.1644, 0.2006]$
$[0.1024, 0.1363, 0.1650]$	$[0.1706, 0.2053, 0.2433]$

TABLE 6.5: In this table we report the set of quantiles obtained from the distribution of the multifractality proxy $\Delta H(1,2)$ on the two tranches T_1 and T_2 for 1000 simulated time series whose hierarchical order $n_{T_1} < n_{T_2}$. Probabilities p's of the model are initialised to take values in $[0.4, 0.6]$. One observes an almost rigid shift towards larger values of the distribution of the observed $\Delta H(1,2)$ when the order increases.

$n_{T_1} > n_{T_2}$	
T_1	T_2
$[Q^{T_1}_{2.5\%}, Q^{T_1}_{50\%}, Q^{T_1}_{97.5\%}]$	$[Q^{T_2}_{2.5\%}, Q^{T_2}_{50\%}, Q^{T_2}_{97.5\%}]$
$[0.1304, 0.1628, 0.1950]$	$[0.1126, 0.1446, 0.1760]$
$[0.1564, 0.1941, 0.2327]$	$[0.1251, 0.1608, 0.1933]$
$[0.2061, 0.2484, 0.2893]$	$[0.1313, 0.1651, 0.1993]$
$[0.1064, 0.1375, 0.1673]$	$[0.0792, 0.1079, 0.1332]$
$[0.0852, 0.1124, 0.1411]$	$[0.0257, 0.0517, 0.0736]$
$[0.1002, 0.1328, 0.1624]$	$[0.0778, 0.1086, 0.1343]$
$[0.1514, 0.1902, 0.2253]$	$[0.1051, 0.1348, 0.1642]$
$[0.2086, 0.2471, 0.2954]$	$[0.0847, 0.1127, 0.1403]$
$[0.1093, 0.1425, 0.1721]$	$[0.0517, 0.0789, 0.1016]$
$[0.1913, 0.2222, 0.2743]$	$[0.1042, 0.1359, 0.1629]$
$[0.1106, 0.1427, 0.1759]$	$[0.0827, 0.1113, 0.1387]$

TABLE 6.6: In this table we report the set of quantiles obtained from the distribution of the multifractality proxy $\Delta H(1,2)$ on the two tranches T_1 and T_2 for 1000 simulated time series whose hierarchical order $n_{T_1} > n_{T_2}$. Probabilities p's of the model are initialised to take values in $[0.4, 0.6]$. One observes an almost rigid shift towards smaller values of the distribution of the observed $\Delta H(1,2)$ when the order decreases.

$n_{T_1} = n_{T_2}$	
T_1	T_2
$[Q^{T_1}_{2.5\%}, Q^{T_1}_{50\%}, Q^{T_1}_{97.5\%}]$	$[Q^{T_2}_{2.5\%}, Q^{T_2}_{50\%}, Q^{T_2}_{97.5\%}]$
$[0.1559, 0.1936, 0.2312]$	$[0.1490, 0.1856, 0.2212]$
$[0.1773, 0.2122, 0.2522]$	$[0.1699, 0.2089, 0.2456]$
$[0.1306, 0.1639, 0.1988]$	$[0.1239, 0.1597, 0.1930]$
$[0.0818, 0.1107, 0.1378]$	$[0.0784, 0.1099, 0.1360]$

TABLE 6.7: In this table we report the set of quantiles obtained from the distribution of the multifractality proxy $\Delta H(1,2)$ on the two tranches T_1 and T_2 for 1000 simulated time series whose hierarchical order $n_{T_1} = n_{T_2}$. Probabilities p's of the model are initialised to take values in $[0.4, 0.6]$. One observes an almost rigid shift towards smaller values of the distribution of the observed $\Delta H(1,2)$s when the order decreases.

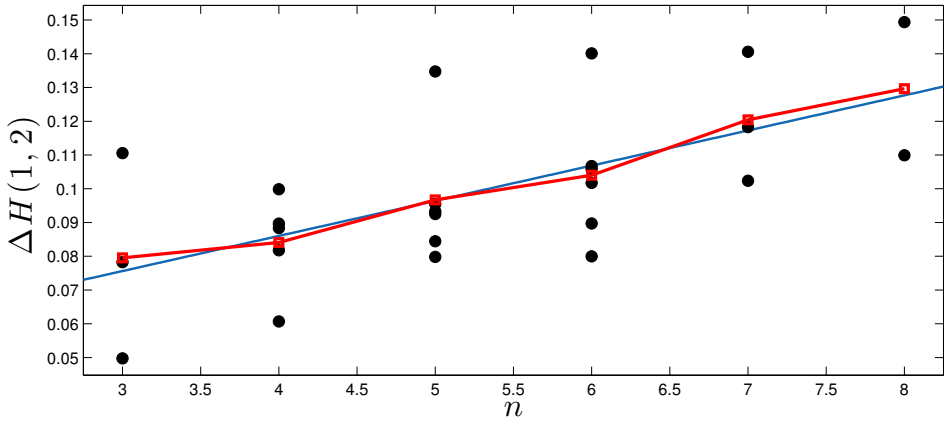

FIGURE 6.17: We plot the average multifractality indicator $\langle \Delta H(1,2) \rangle$ 25 MDHM multivariate time series with the hierarchical structure shown in Figure 6.8 and probabilities randomly drawn in $[0, 1]$.

complexity and risk and gives a novel mechanism for interpreting the source of multifractality in financial markets. We have simulated 25 synthetic MDHM series of time length 4000 using the hierarchy in Figure 6.8 obtained from the data and computed $\Delta H(1,2)$ on each synthetic series. The multifractal properties of the signal are proportional to the order of the hierarchy n. Multifractality is indeed larger for those stocks whose order is larger, as can be seen in Figure 6.17. This fact shows that multifractality is significantly increased by the collective action of many risks, as is the case when the hierarchical order of each $r_{i,t}$ is increased. A more robust quantitative assessment of this property of the model is achieved

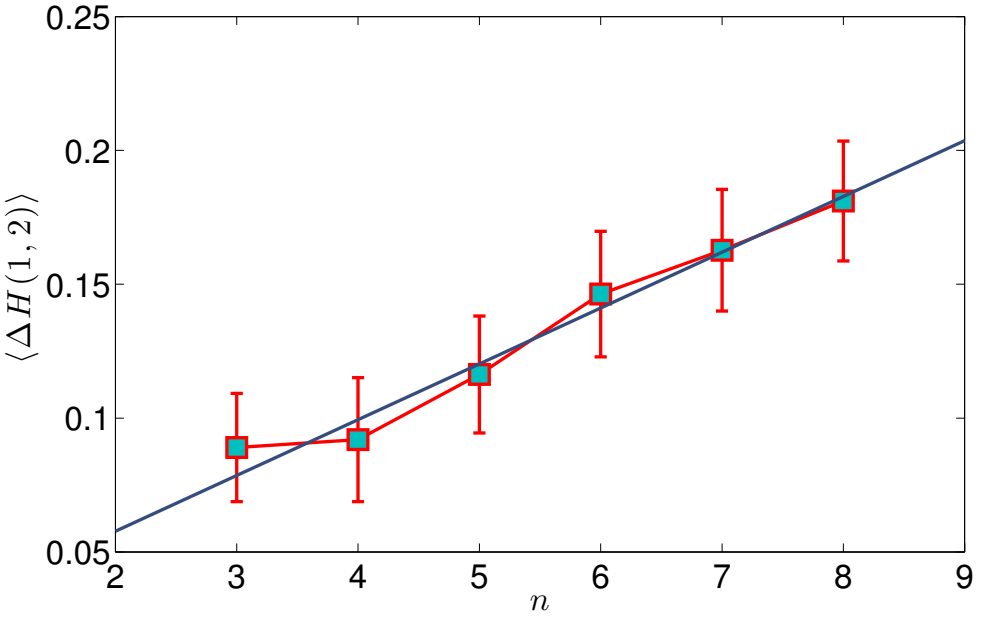

FIGURE 6.18: We plot the average multifractality indicator $\langle \Delta H(1,2) \rangle$ over 1000 independent simulations of DHM multivariate time series with the hierarchical structure shown in Figure 6.8 and probabilities randomly drawn in $[0,1]$.

simulating 1000 independent replications of MDHM processes with the same hierarchical structure and averaging multifractality over the same hierarchical order. The result is shown in Figure 6.18.

The intuition behind this mechanism is that high multifractality of the signals is dominated by large fluctuations rather than by long-range dependence. This is in agreement with the empirical observations reported in many studies [16, 17] and with the analysis performed in Chapter 2 of this thesis.

Note that within the MDHM framework, one can ascribe non-stationarity of multifractality entirely to the the changing distribution tail index α, without having to invoke a breakdown in the autocorrelation structure of the volatility, which would be implied if one would for example assume a time-varying intermittency λ in the MRW [2]. Hence, the factorisation of the hierarchical term from one common volatility factor whose autocorrelation function does not change in time has the advantage of being easier to handle than a model where the autocorrelation function is time-varying.

[2] λ is in fact responsible for both volatility autocorrelation and tail exponent, as seen in Chapter 3.

6.4 Summary

In this chapter we have introduced a model which incorporates explicitly a time varying hierarchical structure of risks into the volatility factor. We have discussed the model construction and the dependence on its key parameters. The presence of the hierarchical factor in the volatility results in a perturbation to the correlation structure of the multivariate time series which is shown to include also the empirical observations not accounted for by a standard multivariate model with one common volatility factor only.

The model presented in this chapter stands out from all other multifractal models in that it ascribes the observed non-stationarity of multifractal properties entirely to the fat-tailed nature of the distribution. Remarkably, a construction based on a hierarchy of common factors proves to be able to introduce a non-trivial dependence between the hierarchical order of each stock and its multifractal properties. This mechanism gives a new interpretation to the latent organisation of financial markets and is the first attempt to explain coherently two aspects considered hitherto completely separate. It is also the first model to merge together the use of network theory tools (the clustering algorithm to detect the hierarchical structure) and multivariate models.

Chapter 7

Conclusions and Outlook

"One thing we have lost, that we had in the past, is a sense of progress, that things are getting better. There is a sense of volatility, but not of progress."

- D.Kanheman -

In this thesis we studied two fundamental aspects of financial time series: multifractality and cross-correlation. We mainly unveiled novel empirical properties that had never been detected before and proposed a model able to reproduce these observations. Main results were the following:

- we introduced the weighted generalised Hurst exponent and investigated the dynamical properties of scaling exponents on companies defaulted or bailed-out after the 2008 credit crisis
- we studied the dynamical multifractal properties of a large set of daily equity prices, showing that the hypothesis of stationarity of the multi-scaling behaviour does not hold globally
- we found interplay between multifractal and cross-correlation hierarchical properties providing, for the first time in the literature, a link between two different aspect of complexity in financial markets

- we proposed a heuristic model able to reproduce the observed stylised facts, which extends a log-normal dynamical multivariate model incorporating the correlation hierarchy explicitly.

We first highlighted in Chapter 2 how scaling exponents can convey information about the stability of the market shares of a firm over turbulent periods. Then, in Chapter 3, we discussed how, if financial time series are assumed to be described by a constant intermittency multifractal model, the proxy of multi-scaling obtained through the wGHE method reveals time variations much larger than those envisaged by the model and therefore the hypothesis of constant multifractality is at odds with empirical observations. We then investigated in Chapter 4 the interdependence between multifractality and cross-correlation hierarchical structure and found that these two apparently different aspects of complexity show a remarkable interplay which does not allow to consider them as completely separate. To capture this new empirical fact we have then proposed in Chapter 6 a new model, built on existing multivariate models which we have extensively studied and discussed in Chapter 5. The model is, to the best of our knowledge, the first dynamical model to include a time varying hierarchy of risks in its structure and it has been shown to be able to reproduce the newly found empirical interdependence between cross-correlation hierarchical structure and multifractality. The model can also account for time varying multifractality and this feature is coherently linked to the dynamical variation of the cross-correlation hierarchical structure.

The research carried out in this thesis may well serve as an empirical basis to investigate further possible coherent underlying mechanisms governing price formation and evolution. The non-stationary behaviour of multifractality is likely to disappear as the return time scale is made coarser, but may reveal even more complex properties in the high-frequency regime. In this regard, a more detailed study across different time scales might reinforce the results found in this thesis and shed more light on which factors affect the non-stationarity observed in financial time series.

The empirical finding discussed in Chapter 4 are also very promising in opening new research avenues. Until now in fact, multifractality and cross-correlation had

always been treated as two separate aspects and no attempt to understand them as part of a coherent framework had been made. The observations reported in this thesis point in the direction of some fundamental mechanism tying the two features together, like that proposed in Chapter 6. Much progress is likely to come from the combination of network theory tools and stochastic modelling techniques, which have been seldom merged together for modelling purposes so far.

Another interesting finding of this thesis is the apparently different behaviour of Asian and Western markets. While both Tokyo and Hong Kong stock exchanges appear to be dominated by one single market mode, NYSE and LSE show much more complex structure, which is well captured by the DBHT hierarchical construction: in loose words one can say that correlation hierarchies in Asian markets are much deeper and narrower, while they are spread horizontally in western markets. These observations go along with multifractality being much more heterogeneous in the American and British market that in the Japanese and Chinese ones. The consequence is that diversification could be preferable in Western markets, where a richer correlation structure allows to hedge the risk more safely and to exploit different segments of the market to make profits.

As far as multifractality is concerned, the analysis performed in this thesis suggests that a *pure* multifractal behaviour is only detectable when the market is not going through periods of jitters. Over major crisis or when large fluctuations kick in, the observed multifractality is dominated by the tails of the unconditional distribution of returns rather than by proper scaling features. For this reason, multifractal models as the MRW need to be modified and the DHM is a possible modification to account for the empirical observations. In this regard, the findings of this thesis strongly support what originally discussed in [77], where the author remarked how the removal of even single data points - corresponding to major financial crashes - could lead to completely different multifractal scenarios. The message is that, although elegant and sound the cascade picture may be, non-linearities and complexity arising during periods of financial turmoil may have a simpler explanation in terms of extreme fluctuations, rather than on multi-scaling patterns.

Other future investigations concern possible extensions or modifications of the

MDHM model introduced in Chapter 6. A more realistic (yet more complicated) picture is the one where the probabilities of the risks are correlated according to the topology of the dendrograms and cascades or feedback mechanisms between different risks are allowed. Another assumption that would not be far-fetched is to assume probabilities to be larger for risks higher in the hierarchical structure and smaller for risks deep down the hierarchy, which should reflect the fact that the more common a risk is, the more likely it is to impact the stocks affected. All these options remain to be explored.

As far as the DHM model is concerned, further tests and additional studies should be performed in order to verify its self-consistency and its ability to reproduce financial market stylised facts. A first important issue concerns the correlation matrix: the empirical covariance matrix in fact is used in our simulation as the matrix of the multivariate factor ϵ and therefore the correlation retrieved from the simulated series - because of the effect of the hierarchical factor - is necessarily different from the input one. This prevents to check the self-consistency of the model unless one successfully finds out a calibration scheme. Since the DHM model has been formulated following a heuristic approach no calibration scheme has been hitherto devised. In order to calibrate the model, what remains to be understood is a viable way to associate the dendrogram nodes with observable risk factors. A reasonable approach could be that of reducing the number of risks associated with the dendrogram nodes, hence considerably simplifying the problem. Using empirical correlations to calibrate the model leaves the problem still over-parametric as the number of parameters to be calibrated is $N(N-1)/2 + N - 1$ against $N(N-1)/2$ entries of the empirical matrix. Hence, one should find a way of getting rid of $N-1$ components, in such a way though that important information is not thrown away. Network theory tools may be helpful in providing topological constraints which have already proved elsewhere to be able to retain relevant information for portfolio strategies [182].

Another test which would be interesting to perform concerns the ability of the model to reproduce correlation between hierarchical order and multifractality as a function of the average depth of the dendrogram. As we observed in the case

of Asian markets, deep and narrow hierarchical structures tend to sap the dependence between $\langle \Delta H(1,2) \rangle$ and n. Therefore, it would be important to assess whether the DHM is able to capture this further feature of real markets, reproducing dependence for highly diversified hierarchical structures (as already confirmed in this thesis) but not for very compact hierarchical structures.

Finally, from the model formulation it is clear that the hierarchy enters the model only by means of the hierarchical order n: it may therefore make sense to investigate whether it is possible to retrieve the same effect of correlation between multifractality and hierarchical order by dropping all the internal parameters but associating instead a single probability value to n. This would enhance the tractability of the model at the same time retaining the extra benefits generated by the hierarchical structure measured from the data.

This is page intentionally left blank

Appendix A

Markov-Switching Multifractal Model

The Markov-Switching Multifractal Model (MMSM) models returns as [87]

$$r_t = \sigma_t \cdot \epsilon_t, \tag{A.1}$$

where, for each t, ϵ_t is a standard normal random variable, i.e. $\epsilon_t \sim \mathcal{N}(0,1)$ and σ_t is a volatility factor that at each time t depends on the product of k components $M_t^{(1)}, M_t^{(1)}, \ldots, M_t^{(k)}$ through

$$\sigma_t^2 = \sigma^2 \prod_{i=1}^{k} M_t^{(i)} \tag{A.2}$$

with σ a constant factor. Each component $M_t^{(i)}$ for $i = 1, \ldots, k$ is renewed at time t with probability γ_i or left unchanged with probability $1-\gamma_i$, where γ_i is specified by [183]

$$\gamma_i = 1 - (1 - \gamma_k)^{(b^i - k)}, \quad i = 1, \ldots, k. \tag{A.3}$$

The parameters can assume values $\gamma_k \in [0, 1]$ and $b \in (1, \infty)$, but a common choice in the literature [61–63] is to set $b = 2$ and $\gamma = 0.5$ to have

$$\gamma_i = 1 - \left(\frac{1}{2}\right)^{(2^i - k)}, \quad i = 1, \ldots, k. \tag{A.4}$$

For the volatility components a popular choice [17, 63] is to consider the Binomial distribution, that is $M_t^{(i)} \sim \{m_0, 2 - m_0\}$, with $1 \leq m_0 < 2$.

Appendix B

Relationship between Hurst exponent and tail index

In order to better understand the link between scaling exponents and tail exponents, let us consider the simple ideal case where the underlying process is a random walk with $x_t = x_{t-1} + \eta_{t-1}$ where $x_t = \ln(P_t)$. In this case, for an arbitrary τ, the log-returns $r_{t,\tau} = x_{t+\tau} - x_t$ can be written as a sum of $n = \tau$ random variables:

$$r_{t,\tau} = \sum_{s=0}^{\tau-1} \eta_{s+t}. \tag{B.1}$$

If the η_t are iid, the Central Limit Theorem applies to the $r_{t,\tau}$ and there are two cases: (1) the probability distribution function of η_t has finite variance and therefore the distribution of $r_{t,\tau}$ converges to a normal distribution for large τ; (2) the variance is not defined and the asymptotic distribution of the $r_{t,\tau}$ converges to a Levy Stable distribution. For distributions well approximated by power-law functions in the tail region, the parameter that distinguishes between these two cases is the tail index α. Namely $\alpha \geq 2$ leads to normal distributions, while $\alpha < 2$ leads to Levy Stable distributions. Moreover, given that $r_{t,\tau}$ is a sum of random

variables and given that both cases (1) and (2) lead to stable distributions[1] , the probability distribution $p_\tau(r)$, of the log-returns must scale with τ as [1, 27]

$$p_\tau(r) = \begin{cases} \frac{1}{\tau^{1/\alpha}} p\left(\frac{r}{\tau^{1/\alpha}}\right) & \text{if } \alpha < 2 \\ \\ \frac{1}{\tau^{1/2}} p\left(\frac{r}{\tau^{1/2}}\right) & \text{if } \alpha \geq 2 \, . \end{cases} \tag{B.3}$$

Accordingly, the q-moments scale as

$$E(|r_{t,\tau}|^q) = \begin{cases} \tau^{q/\alpha} E(|r_{t,1}|^q) & \text{if } \alpha < 2 \\ \\ \tau^{q/2} E(|r_{t,1}|^q) & \text{if } \alpha \geq 2 \, . \end{cases} \tag{B.4}$$

Here $E(...)$ denotes the expectation value. Finally, if we restrict to the class of self-affine processes, i.e. those processes $x(t)$ where the probability distribution of $\{x(ct)\}$ is equal to the probability of $\{c^H x(t)\}$, for any positive c, and we consider stationary increments, the q-moments must scale as

$$E(|r_{t,\tau}|^q) = c(q)\tau^{qH} \, . \tag{B.5}$$

By comparing Equation B.4 with Equation B.5 we get

$$H = \begin{cases} 1/\alpha & \text{if } \alpha < 2 \\ \\ 1/2 & \text{if } \alpha \geq 2 \, . \end{cases} \tag{B.6}$$

Equation B.5 holds also for the moments computed using the weighted average, by substituting H with H^w and the expectation values $E(\dots)$ with weighted averages. Processes with the property in Equation B.5 are deemed uniscaling. For $\alpha \geq 2$

[1] A distribution is stable if and only if, for any $n > 1$, the distribution of $y = x_1 + x_2 + \cdots + x_n$ is equal to the distribution of $n^{1/\alpha}x + d$, with $d \in \mathbb{R}$. This implies

$$p_n(y) = \frac{1}{n^{1/\alpha}} p\left(\frac{y-d}{n^{1/\alpha}}\right) \tag{B.2}$$

where $p_n(y)$ is the aggregate distribution of the sum of the i.i.d. variables and $p(x)$ is the distribution of the x_i.

we retrieve $H \sim 0.5$ and the processes scales as a Brownian motion. For $\alpha < 2$ instead the scaling exponent increases with the reciprocal of α and therefore small α's can cause H to increase. Let us here stress that the result in Equation B.6 is only valid for a random-walk type iid process with defined noise distribution and it is well known that financial time series cannot be described within this framework. However, Equation B.6 is a valuable reference which can be used as a tool to compare the relation between the tail exponent and the Hurst exponent in more complex signals.

Appendix C

GARCH(p,q) models

GARCH models were introduced in the '80s with the aim of capturing the heteroskedastic nature of volatility. Since then, many generalisations and modifications have been investigated by econometricians and statisticians.
In the GARCH(p, q) frameworks, returns are modelled as [86]

$$\begin{aligned} r_t &= \sigma_t \epsilon_t, \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{p} \alpha_i r_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \end{aligned} \tag{C.1}$$

where $\alpha_0 > 0$, $\alpha_i \geq 0 \ \forall i = 1, \ldots, p$ and $\beta_j \geq 0 \ \forall j = 1, \ldots, q$ are all constants. In practice only low-order GARCH models are used and the GARCH(1,1) reads

$$\sigma_t^2 = \alpha_0 + (\alpha_1 \epsilon_{t-1}^2 + \beta_1)\sigma_{t-1}^2. \tag{C.2}$$

One can show that this model's volatility and square returns autocorrelation functions decay exponentially with the lag ℓ [27] and therefore are not capable of capturing the long-range dependence of financial time series. In general, one exponential time scale for the autocorrelation function corresponds to short range dependence (the autocorrelation function is integrable), while the superposition of many exponential time scales can improve the situation significantly, as shown in [89]. Examples of a realisation of GARCH(1,1) process are given in Figure C.1.

The literature on GARCH related models is very reach and includes contributions

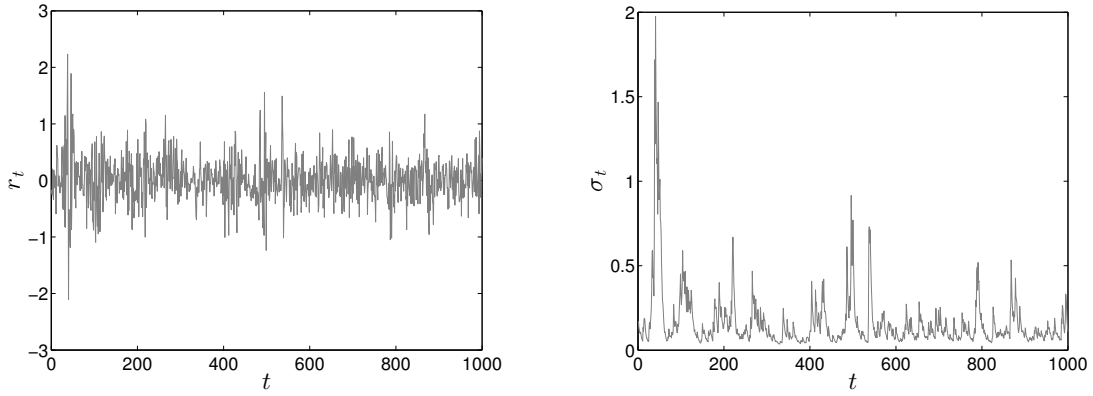

FIGURE C.1: Returns (left) and conditional variance (right) for a GARCH(1,1) process simulated with parameters $\alpha_0 = 0.01, \alpha_1 = 0.25$ and $\beta_1 = 0.7$.

from both the econometric community (see Chapter 4 in [132]) and Econophysics [137].

Appendix D

Deterministic Bubble Hierarchical Tree (DBHT) clustering technique

In this Appendix we discuss the DBHT clustering algorithm, which has been employed to determine the hierarchical structure of our datasets throughout this thesis. As already pointed out in Chapter 4, this algorithm is preferred to other existing clustering techniques for four main reasons: 1) it has been shown in [22] that it outperforms other clustering methods in recognising through cluster organisation existing partitions present in the data, 2) it naturally provides both an inter and intra-cluster organisation without the need to resort to further arbitrary choices, 3) it doesn't necessarily imply that internal junction on the dendrogram be ranked according to increasing or decreasing correlation, making it possible to associate the output dendrograms to the hierarchical model presented in Chapter 6 and 4) it is a deterministic algorithm.

In explaining the clustering method we follow the original formulation presented in [22]. The DBHT uses the topological structure of the PMFG to produce a hierarchical structure of clusters. Let us denote the weights of the edges uv in the PMFG as $w_{u,v}$. The weights depend on a similarity measure and therefore larger weights correspond to stronger similarity. We also associate to each edge

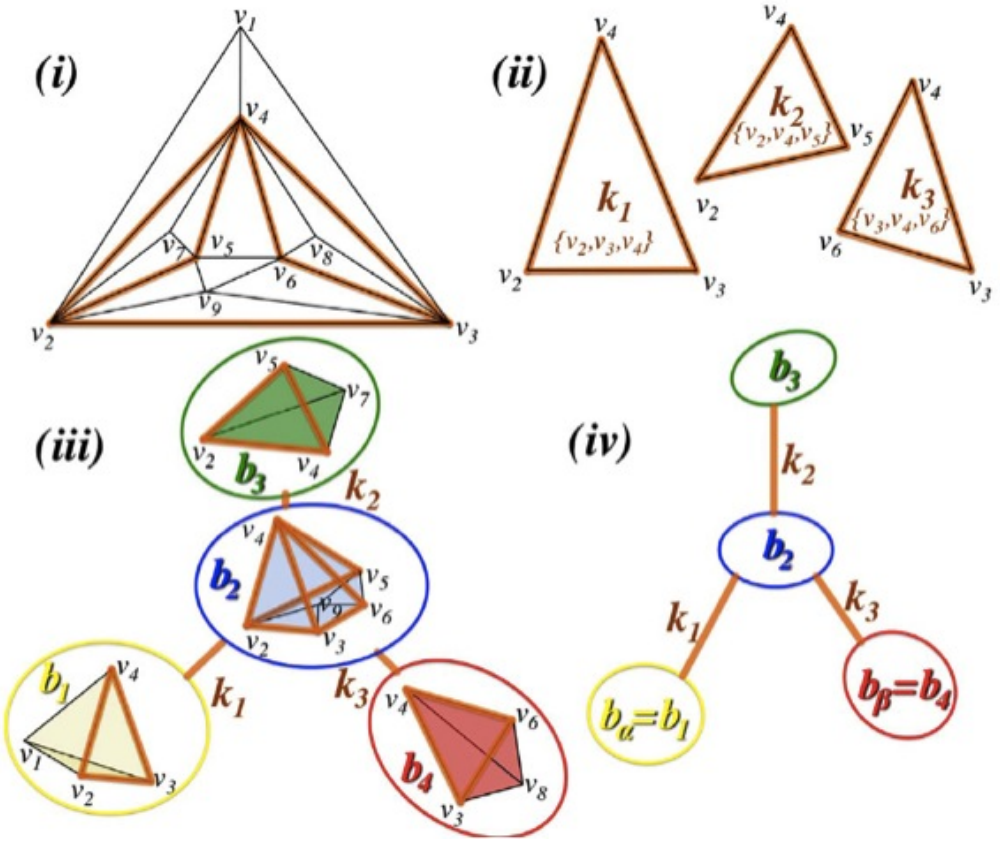

FIGURE D.1: (i) An example of the PMFG with nine vertices, $V(G) = \{v_i\}_{i=1,\ldots,9}$ and containing three separating 3-cliques: k_1, k_2 and k_3. (ii) The separating 3-cliques have vertex sets: $V(k_1) = \{v_2, v_3, v_4\}$, $V(k_2) = \{v_2, v_4, v_5\}$ and $V(k_1) = \{v_3, v_4, v_6\}$. (iii) The separating 3-cliques identity four planar subgraphs called *bubbles*: b_1, b_2, b_3 and b_4. (iv) The graph can be viewed as a *bubble tree* made of four bubbles connected through separating 3-cliques. Figure taken from reference [22].

uv a non-negative dissimilarity measure $d_{u,v}$, so that the PMFG graph can be conveniently expressed as $G(V, E, W, D)$, where V is the set of vertices, E the set of edges, W the set of weights and D the set of dissimilarities between each pair of vertices. The DBHT method then looks for separating 3-cliques in the PMFG. Each separating 3-clique k_p divides the graph G into two disconnected parts, the interior G_p^{in} and the exterior G_p^{ex}, that are joined by the clique itself. The presence of cliques within cliques naturally provides a hierarchy and the subdivision into 3-cliques can be carried out until the entire graph is considered. This results in a set of planar graphs that are dubbed *bubbles*, connected to each other via the separating 3-cliques, and that can be visualised as a tree H_b, as shown in Figure D.1.

The next step of the methods associates direction to each edge in H_b by comparing the sum over the weights of the edges in the PMFG connecting the 3-clique k_p with the two bubbles, obtaining a directed bubble tree $\overrightarrow{H}_b$. It is then possible to distinguish among three types of bubbles in $\overrightarrow{H}_b$: (i) converging bubbles, where all edges are incoming to the bubble, (ii) diverging bubbles, where all edges are outgoing and (iii) passage bubbles, where both types of edges are present. As converging bubbles are somehow special being the end points of directional paths that follow the strongest connections, they are considered as the centres of the clusters. Note that no explicit ranking in the strengths of correlation is introduced (as in other clustering methods). Any bubble b_i connected by a directed path in $\overrightarrow{H}_b$ to a converging bubble b_α belongs to cluster α. By construction, bubbles in cluster α form a subtree which has only one converging bubble b_α and all edges are directed toward b_α.

Then, in order to obtain a clustering of the vertex set $V(G)$, we assign to each vertex v the cluster membership of the bubble containing it as shown in Figure D.2. This is done by considering first vertices belonging to converging bubbles and then all other vertices. Vertices belonging to only one converging bubble are simply assigned to it, whereas for vertices belonging to more than one converging bubble (e.g. vertices v_2 and v_3 in Figure D.2) the strength of the attachment is evaluated as

$$\chi(v, b_\alpha) = \frac{\sum_{u \in V(B_\alpha)} A_G(v,u)}{3(|V(b_\alpha)| - 2)}, \tag{D.1}$$

with $A_G(v,u)$ being the uv-entry of the adjacency matrix associated to the PMFG, and the vertex is assigned to the bubble with largest strength. Then the set of all remaining vertices is considered (e.g. vertices v_5, v_7 and v_9 in Figure D.2). A vertex v may in principle belong to more than one subtree $\overrightarrow{h}_\alpha, \overrightarrow{h}_\beta, \ldots$ and, in this case, it is assigned to the converging bubble that has the minimum average shortest path distance

$$\bar{L}(v, \alpha) = \text{mean}\{l(v,u) | u \in V^0(\alpha) \vee v \in V(\overrightarrow{h}_\alpha)\}, \tag{D.2}$$

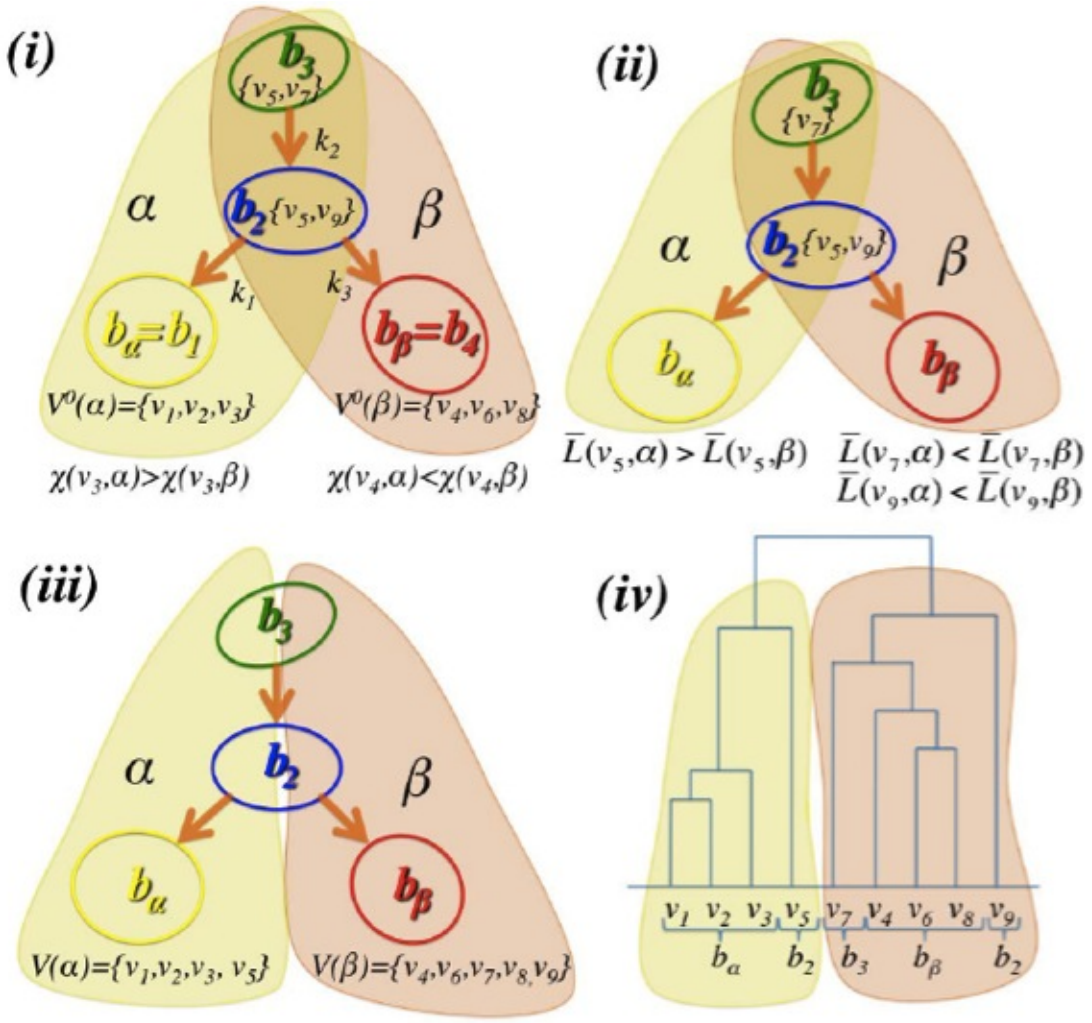

FIGURE D.2: (i) Construction of the directed bubble tree where directions are given to the 3-cliques according to their largest weights. In this example there are two converging bubbles: $b_\alpha = b_1$ and $b_\beta = b_4$. A unique set of vertices can be associated to each of the converging bubbles where vertices shared by both two bubbles (i.e. v_3 and v_4) are assigned according to the largest strength χ. (ii) All other non-assigned vertices are associated to the cluster with minimum $\bar{L}$. (iii) The vertex set is uniquely divided into two clusters respectively associated to the two converging bubbles: $V(\alpha) = \{v_1, v_2, v_3, v_5\}$ and $V_\beta = \{v_4, v_6, v_7, v_8, v_9\}$. (iv) The hierarchical organisation and is represented via a dendrogram. Figure taken from reference [22].

where $V^0(\alpha)$ is the unique set of vertices in the converging bubble α assigned in the first step and $l(v,u)$ is the shortest path distance on G from v to u, i.e. the shortest sum of distances $d_{r,s}$ between v and u. This completes the partition of the vertex set into discrete clusters.

Once the partition into clusters has been identified, one can proceed to investigate how these clusters are internally structured and what is the hierarchical organisation of different clusters. This is achieved by building the hierarchy at three separate levels:

1. **Intra-bubble Hierarchy:** first, each vertex $v \in V(\alpha)$ is assigned to a buble b_i in the subtree $\overrightarrow{h}_\alpha$, while vertices in the converging bubble have been already assigned to to the sets $V^0(\alpha)$. For all remaining vertices, those belonging to only one bubble are assigned to such bubble, whereas those belonging to more than one bubble are assigned to the bubble that maximises $\chi(v,b_i)$, as shown in Figure D.2. On each unique vertex set $V^\alpha(b_i)$ consisting of vertices for each bubble b_i in $\overrightarrow{h}_\alpha$, one can therefore perform a complete linkage procedure by using $l(u,v)$ as the distance matrix.

2. **Intra-cluster Hierarchy:** the complete linkage procedure is performed between bubbles in $\overrightarrow{h}_\alpha$ by using the distance matrix

$$d^I(b_i,b_j) = \max\{l(u,v)|u \in V^\alpha(b_i) \vee v \in V^\alpha(b_j)\}. \tag{D.3}$$

3. **Inter-cluster Hierarchy:** the complete linkage procedure is performed on the distance matrix

$$d^{II}(\alpha,\beta) = \max\{l(u,v)|u \in V(\alpha) \vee v \in V(\beta)\}. \tag{D.4}$$

In this way one obtain a linkage procedure that starts from the discrete clusters and at higher level joins different clusters into superclusters, whereas it splits the clusters at lower levels into a hierarchy of bubbles and the bubbles in turn into a hierarchy of elements. The output of this method is the dendrogram that is used to retrieve the hierarchical order n used throughout the thesis.

Appendix E

$\Delta H(1,2)$ vs n on Sub-clusters

In this Appendix we show plots of $\Delta H(1,2)$ vs n for more clusters and sub-clusters. As for clusters shown in Chapter 4 we find an average positive dependence between multifractality and hierarchical order, confirming that stocks with higher degree of multifractality are also those deeper in the hierarchy of correlations.

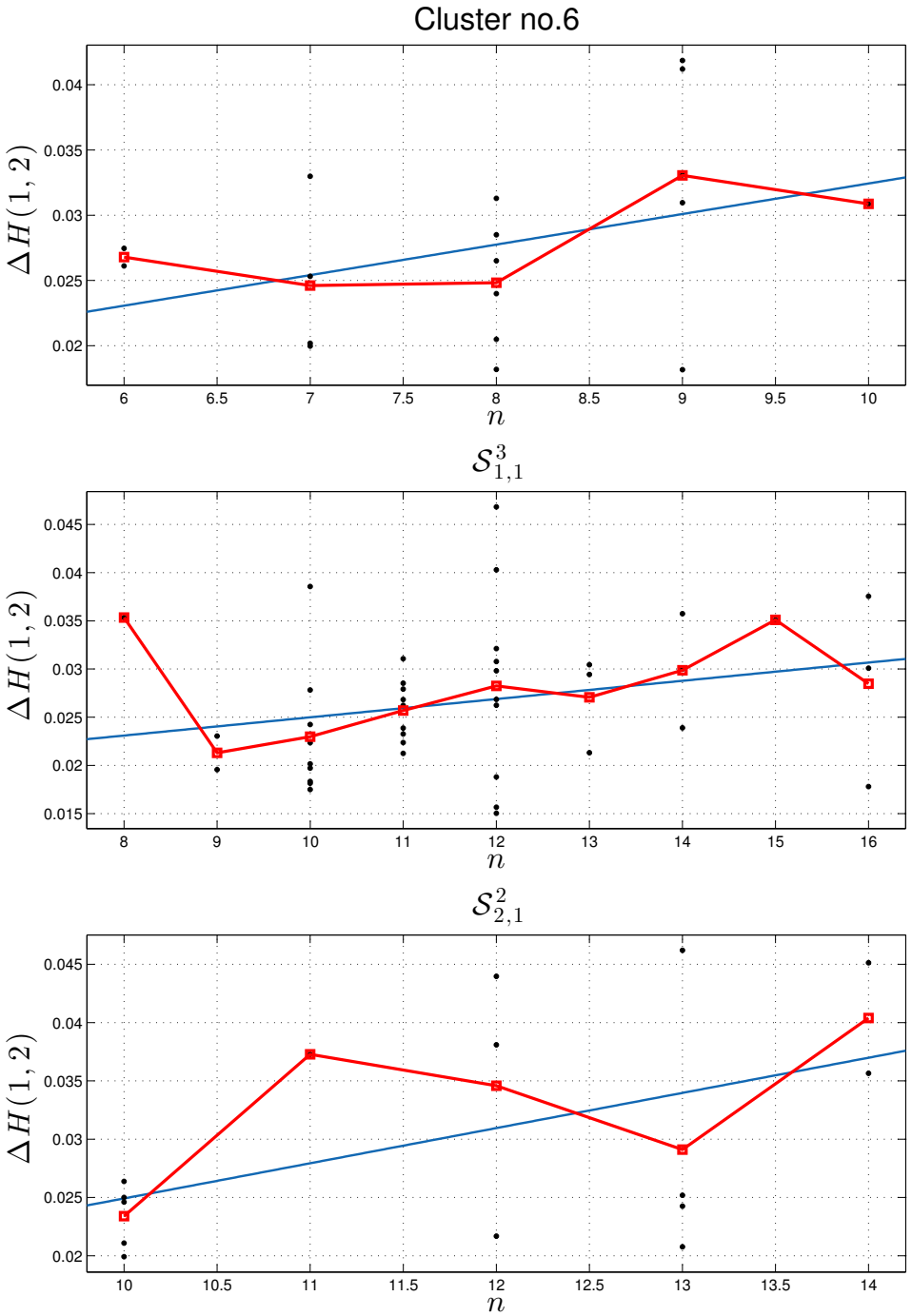

FIGURE E.1: $\Delta H(1,2)$ vs n on Sub-clusters. Black dots are $\Delta H(1,2)$ against the hierarchical order n for stocks in cluster no.6 (top) and in the sub clusters $\mathcal{S}^3_{1,1}$ (middle) and $\mathcal{S}^2_{2,1}$ (bottom). In all three plots the blue line is the best fit of the dots, while the red squares are the averages of $\Delta H(1,2)$ for each fixed order.

Appendix F

Generalities on Multivariate Statistics

Multivariate models attempt to give a sound statistical description of the dependency structure observed in stock returns. The main object of interest is thus the joint distribution of stock returns, which encodes the probability of observing common movements of a certain size in a multivariate set of stocks. Let us consider a multivariate vector of stock returns, which we will denote as $\mathbf{X}_t = (X_{1,t}, \ldots, X_{N,t})$. When there is not explicit reference to time, we will drop the subscript t in the following for better legibility. The joint distribution function of stock returns is [132]

$$F_{\mathbf{X}}(\mathbf{x}) = F_{\mathbf{X}}(x_1, \ldots, x_N) = \mathbb{P}(\mathbf{X} \leq \mathbf{x}) = \mathbb{P}(X_1 \leq x_1, \ldots, X_N \leq x_N) \tag{F.1}$$

where we denote by $\mathbb{P}\{\cdot\}$ the probability. The joint distribution (F.1) can be obtained, provided that $\mathbf{X}_t$ is absolutely continuous, from the joint density $f_{\mathbf{r}}$ as

$$F_{\mathbf{X}}(\mathbf{x}) = \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_N} f_{\mathbf{X}}(u_1, \ldots, u_N) du_1 \ldots du_N. \tag{F.2}$$

From (F.1) one can recover the marginals, i.e. the univariate distributions of each of the single variables X_s for $s = 1, \ldots, N$ integrating out all other $N-1$ variables

$$F_{X_i}(x_i) = \mathbb{P}(X_i \leq x_i) = F_{\mathbf{r}}(\infty, \ldots, x_i, \ldots, \infty). \tag{F.3}$$

Independence of $X_1, \ldots, X_N$ corresponds to the distribution function being equal to the product of the N marginals (F.3), i.e.

$$F_{\mathbf{X}}(\mathbf{x}) = \prod_{s=1}^{N} F_{X_s}(x_s). \tag{F.4}$$

Restricting our attention on two variables X_i and X_j, these are in fact independent if the knowledge of one of the two doesn't add any information to the knowledge of the other. Since the joint density of (X_i, X_j) can be factorized as

$$f_{X_i,X_j}(u_i, u_j) = f_{X_i|X_j}(u_i) f_{X_j}(u_j) \tag{F.5}$$

with $f_{X_i|X_j}(u_i)$ being the conditional density of X_i given the knowledge of X_j and (u_i, u_j) an arbitrary pair in the support of (X_i, X_j), one sees that if the two r.v.'s are independent we simply have $f_{X_i|X_j}(u_i) = f_{X_i}(u_i)$ and hence

$$f_{X_i,X_j}(u_i, u_j) = f_{X_i}(u_i) f_{X_j}(u_j). \tag{F.6}$$

Generalising to the N variables and integrating the joint density as in Equation (*F.2*) we obtain Equation (F.4).

A crucial role in the parametrisation of the joint distribution is played by the covariance matrix, which we will denote by $\mathbf{\Sigma}$ and is given by

$$\mathbf{\Sigma} = \mathrm{Cov}(\mathbf{X}) := \mathbb{E}\left[(\mathbf{X} - \mathbb{E}(\mathbf{X}))(\mathbf{X} - \mathbb{E}(\mathbf{X}))^T\right] \tag{F.7}$$

with the expectation operator $\mathbb{E}(\cdot)$ acting componentwise on matrices and yielding the following entries for the covariance matrix

$$\Sigma_{ij} = \mathrm{Cov}(X_i, X_j) = \mathbb{E}(X_i X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j). \tag{F.8}$$

Note that the N diagonal elements of Σ_{ij} are the variances of X_i for $i = 1, \ldots, N$

$$\Sigma_{ii} = \Sigma_i = \mathbb{E}(X_i^2) - \mathbb{E}(X_i)^2. \tag{F.9}$$

Normalising the covariance with the square roots of the variances, one can then define the correlation matrix $\boldsymbol{\rho}$ with entries

$$\rho_{ij} = \frac{\Sigma_{ij}}{\sqrt{\Sigma_i}\sqrt{\Sigma_j}}. \tag{F.10}$$

The correlation coefficient (F.10) accounts for linear relationship between X_i and X_j. Suppose in fact that the two variables are related by an arbitrary linear relation $X_j = \beta\, X_i + \eta$, where η is a residual uncorrelated with X_i and β is a coefficient of linear regression. Since $\Sigma_{ij} = \beta\, \Sigma_i$, one has

$$\rho_{ij} = \beta \sqrt{\frac{\Sigma_i}{\Sigma_j}} \tag{F.11}$$

which shows that the ρ_{ij} is indeed proportional to the linear regression coefficient β.

The estimation of (F.10) is usually carried out via the Pearson's estimator [184–186], defined as

$$\hat{\rho}_{ij} = \frac{\sum_{k=1}^{T} (X_i^k - \langle X_i \rangle) \cdot (X_j^k - \langle X_j \rangle)}{\sqrt{\sum_{i=1}^{T} (X_i^k - \langle X_i \rangle)^2 \cdot \sum_{i=1}^{T} (X_j^k - \langle X_j \rangle)^2}} \tag{F.12}$$

where

$$\langle X_i \rangle = \frac{1}{T} \sum_{k=1}^{T} X_i^k, \qquad \langle X_j \rangle = \frac{1}{T} \sum_{k=1}^{T} X_j^k \tag{F.13}$$

are the sample averages. In spite of its popularity among practitioners and academics, linear correlation is know to suffer from several shortcomings, as was first systematically shown in [127]. The drawbacks in using linear correlation as a measure of dependency between financial time series can be summarised as follows:

- Both X_i and X_j must have finite variances otherwise ρ_{ij} is not defined. This is a very delicate issue as the distribution of the two variables is usually unknown and the empirical estimates do not necessarily correspond to the expectation values and can thus be misleading. The situation is particularly tricky when dealing with heavy-tailed distributions, which is very common in finance.

- Two independent random variables are uncorrelated but the converse doesn't hold. As an example, let us take as the two random variables $r_i \sim \mathcal{N}(0,1)$ and $X_j = X_i^2$ (easy example of non linear dependence). The linear correlation of these two variables clearly vanishes as both $\mathbb{E}[X_i^3]$ and $\mathbb{E}[X_i]$ are zero in spite the dependence between X_i and X_j being very strong. The only case where the notions of independence and un-correlatedness coincide is that of bivariate normal random variables. When only the marginals are normal but the joint distribution is not, correlation is not enough to quantify the entire dependency structure.

- The Pearson estimator (F.12) performs well only when the empirical distributions of X_i and X_j are not fat-tailed. More precisely, the probability of observing events larger than x_i must decay faster than

$$\mathbb{P}(X_i > x_i) \sim x^{-4}. \tag{F.14}$$

  As discussed in Chapter 2, financial time series can show tail exponents even smaller than 2, in which cases the reliability of (F.12) is seriously jeopardised.

- Linear correlation has the further deficiency of not being invariant under non-linear strictly increasing transformations $T : \mathbb{R} \to \mathbb{R}$, i.e.

$$\rho(T(X_i), T(X_j)) \neq \rho(X_i, X_j) \tag{F.15}$$

  a serious pitfall for a measure of dependence, which should not change if the arguments are rescaled.

Aside from the shortcomings listed, linear correlation has the advantage of being straightforward to calculate, involving only the computation of variances and covariance. Other more advanced measures of dependence require more involved tools and may present other problems [147]. Another reason for its widespread use is that the correlation matrix plays a key role as a dependency measure in multivariate Gaussian distributions and in the more general class of elliptical distributions [23, 187], of great relevance in quantitative finance and which we are going to discuss in the next paragraph.

Appendix G

Derivation of Equation (5.17)

For $\mathbf{\Sigma}$ a positive definite matrix, the joint density of $\mathbf{X}$ is given by [132, 148]

$$f(\mathbf{x}) = \int f_{\mathbf{X}|\sigma}(\mathbf{x}|\sigma) dP(\sigma), \tag{G.1}$$

where $f_{\mathbf{X}|\sigma}$ denotes the conditional density of $\mathbf{X}$ given σ and

$$P(\sigma) = \frac{1}{\sqrt{2\pi}\sigma\omega_0} \exp\left(-\frac{\log^2 \sigma}{2\omega_0^2}\right). \tag{G.2}$$

Expanding Equation (G.1) we have

$$f(\mathbf{x}) = \int \frac{\sigma^{-N/2}}{(2\pi)^{N/2}|\mathbf{\Sigma}|^{1/2}} \exp\left(\frac{(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}{2\sigma}\right) dP(\sigma) \tag{G.3}$$

and since $dP(\sigma) = p(\sigma)d\sigma$ with $p(\sigma)$ the pdf of σ, we can write

$$f(\mathbf{x}) = \int \frac{\sigma^{-N/2}}{(2\pi)^{N/2}|\mathbf{\Sigma}|^{1/2}} \exp\left(\frac{(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}{2\sigma}\right) \frac{1}{\sqrt{2\pi}\sigma\omega_0} \exp\left(-\frac{\log^2 \sigma}{2\omega_0^2}\right) d\sigma \tag{G.4}$$

Taking out the constant terms and noting that $\exp(-\log^2 \sigma/(2\omega_0^2)) = \sigma^{-\log \sigma/(2\omega_0^2)}$, we finally get

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{N/2}|\mathbf{\Sigma}|^{1/2}} \frac{1}{\sqrt{2\pi}\omega_0} \int_0^\infty \sigma^{-(\frac{\log \sigma}{2\omega_0^2} + \frac{N}{2} + 1)} \exp\left(\frac{(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}{2\sigma}\right) d\sigma. \tag{G.5}$$

Appendix H

Correlation when $\Gamma_i = \Gamma_j$

When $\Gamma_i = \Gamma_j$ one has $\Gamma_i \setminus \Gamma_j = \emptyset$ and $\Gamma_i \cap \Gamma_j = \Gamma_i = \Gamma_j$. It follows that

$$\mathcal{F}_{ij}(\mathbf{p}; \Gamma_i, \Gamma_j) = \frac{\prod\limits_{x:a_x \in \Gamma_i} \zeta_2(p_x)}{\left(\prod\limits_{m=1}^{n_i} \zeta_2(p_m) \prod\limits_{q=1}^{n_i} \zeta_2(p_m) \right)^{1/2}} = \frac{\prod\limits_{q=1}^{n_i} \zeta_2(p_q)}{\prod\limits_{m=1}^{n_i} \zeta_2(p_m)} = 1 \tag{H.1}$$

and therefore $\rho_{ij} = \mathrm{Corr}(\epsilon_i \epsilon_j)$.

Analogously, the limits $p_m \to 1$ and $p_m \to 0$, $\forall m = 1, \ldots, N-1$ give $\rho_{ij} = \mathrm{Corr}(\epsilon_i \epsilon_j)$. Recalling the expressions for $\zeta_1(p)$ and $\zeta_2(p)$, we have $\lim_{p\to 0} \zeta_1(p) = \lim_{p\to 0} \zeta_2(p) = 1$ and $\lim_{p\to 1} \zeta_1(p) = e$, $\lim_{p\to 1} \zeta_2(p) = e^2$, which implies

$$\lim_{p_m \to 0, \forall m=1,\ldots,N-1} \mathcal{F}_{ij}(\mathbf{p}; \Gamma_i, \Gamma_j) = 1, \qquad \lim_{p_m \to 1, \forall m=1,\ldots,N-1} \mathcal{F}_{ij}(\mathbf{p}; \Gamma_i, \Gamma_j) = 1. \tag{H.2}$$

Appendix I

Selection of 25 stocks

We chose the following 25 stocks from the set (we report the tickers in parenthesis) : (1) First Horizon National Corp (FHN), (2) Abbot Laboratories (ABT), (3) Adobe Systems (ADBE), (4) Advanced Micro Devices INC (AMD), (5) Aflac INC (AFL), (6) Air Products and Chemicals INC (APD), (7) Allergan INC/United States (AGN), (8) Honeywell International/United States (HON), (9) The Allstate Corporation (ALL), (10) Altera Corp (ALTR), (11) Alcoa INC (AA), (12) Barrick Gold Corp (ABX), (13) American Elettric Power CO INC (AEP), (14) American Express Corp (AXP), (15) American Greetings (AM), (16) Amgen INC (AMGN), (17) Hess Corp (HES), (18) American International Group INC (AIG), (19) Anadarko Petroleum Corp (APG), (20) Analog Devices (ADI), (21) Aon Corp (AOC), (22) Apache Corp (APA), (23) Apple INC (AAPL), (24) Applied Materials INC (AMAT), (25) Archer-Daniels-Midlan Co(ADM). The choice was arbitrary and also other choices have been explored without revealing any blatant bias or evident difference.

Appendix J

Hierarchies $\mathcal{H}_{T_1}$ and $\mathcal{H}_{T_2}$

We show in this appendix the dendrograms obtained via DBHT clustering on the set of 25 stocks in Appendix I. These hierarchies of cross-correlations are the ones used to simulate the two-period MDHM in Chapter 6. We remark that, being the DBHT clustering a deterministic algorithm, the hierarchy is not affected by any randomness and performing the clustering many times returns the same result.

	n			n	
Stock	T_1	T_2	Stock	T_1	T_2
1	6	5	14	5	6
2	7	6	15	5	4
3	4	8	16	7	5
4	2	5	17	3	4
5	10	6	18	10	4
6	5	7	19	5	3
7	7	7	20	4	7
8	5	4	21	9	5
9	8	8	22	5	4
10	5	7	23	4	4
11	5	8	24	5	6
12	4	2	25	5	8
13	6	6			

TABLE J.1: In this table we report the hierarchical order n on the two time windows T_1 and T_2 for all 25 stocks. Stocks whose label is coloured in red are those with $n_{T_1} > n_{T_2}$, stocks coloured in blue are those with $n_{T_1} < n_{T_2}$ and those coloured in green have $n_{T_1} = n_{T_2}$.

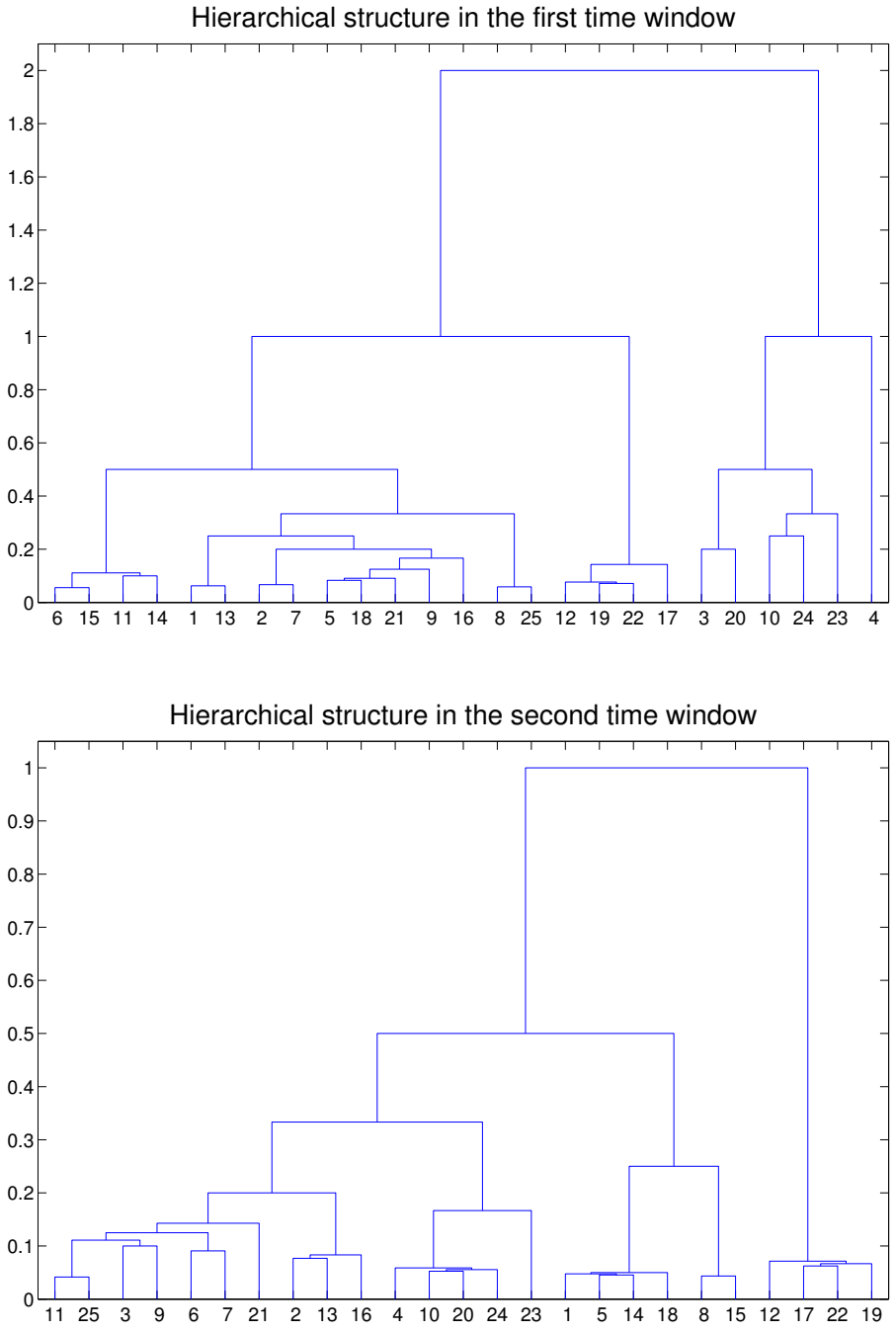

FIGURE J.1: (Top) Hierarchical structure found applying DBHT on the set of 25 stocks in the period 1/1/1995 to 31/12/2004. The hierarchical order $n_{\mathcal{H}_1}$ ranges from 3 to 10. (Bottom) Hierarchical structure found applying DBHT on the set of 25 stocks in the period 31/12/2004 to 22/10/2012. The hierarchical order $n_{\mathcal{H}_2}$ ranges from 2 to 8.

Appendix K

Time-varying quantiles for different probabilities

In this Appendix we report results for time-varying $\Delta H(1,2)$ quantiles obtained on MDHM time series simulated with parameters uniformly drawn in different ranges of values. Specifically, Tables K.1, K.2 and K.3 reports values of the $\Delta H(1,2)$ quantiles for $p \in [0.6, 0.8]$ in the three cases, respectively $n_{T_1} < n_{T_2}$, $n_{T_1} > n_{T_2}$ and $n_{T_1} = n_{T_2}$, whereas tables K.4, K.5 and K.6 show quantiles for $p \in [0.2, 0.4]$.

$n_{T_1} < n_{T_2}$	
T_1	T_2
$[Q^{T_1}_{2.5\%}, Q^{T_1}_{50\%}, Q^{T_1}_{97.5\%}]$	$[Q^{T_2}_{2.5\%}, Q^{T_2}_{50\%}, Q^{T_2}_{97.5\%}]$
$[0.0617, 0.0883, 0.1140]$	$[0.1271, 0.1583, 0.1893]$
$[0.0735, 0.1010, 0.1256]$	$[0.1126, 0.1449, 0.1710]$
$[0.0803, 0.1088, 0.1339]$	$[0.1136, 0.1438, 0.1728]$
$[0.0765, 0.1020, 0.1269]$	$[0.1262, 0.1583, 0.1877]$
$[0.0705, 0.1022, 0.1273]$	$[0.0911, 0.1183, 0.1453]$
$[0.0294, 0.0533, 0.0755]$	$[0.0474, 0.0762, 0.0985]$
$[0.0620, 0.0882, 0.1126]$	$[0.1145, 0.1435, 0.1735]$
$[0.0808, 0.1096, 0.1348]$	$[0.1013, 0.1311, 0.1607]$
$[0.0676, 0.0953, 0.1184]$	$[0.1259, 0.1593, 0.1886]$

TABLE K.1: In this table we report the set of quantiles obtained from the distribution of the multifractality proxy $\Delta H(1,2)$ on the two tranches T_1 and T_2 for 1000 simulated time series whose hierarchical order $n_{T_1} < n_{T_2}$. Probabilities p's of the model are initialised to take values in $[0.6, 0.8]$.

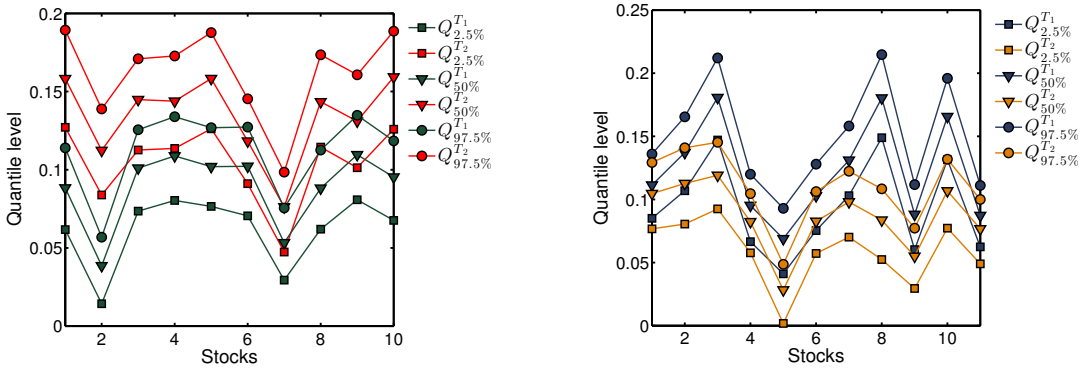

FIGURE K.1: (Top) The plot shows the q-quantiles $Q_q^{T_1}$ and $Q_q^{T_2}$ for $q = \{2.5\%, 50\%, 97.5\%\}$ for the 10 DHM simulated time series whose hierarchical order $n_{T_1} < n_{T_2}$. Red lines correspond to p-quantiles in T_1 while black lines to p-quantiles in T_2. (Bottom) The q-quantiles $Q_q^{T_1}$ and $Q_q^{T_2}$ for $q = \{2.5\%, 50\%, 97.5\%\}$ for the 11 DHM simulated time series whose hierarchical order $n_{T_1} > n_{T_2}$. Blue lines correspond to q-quantiles in T_1 while orange lines to q-quantiles in T_2. The probabilities of the risks are initialised to values ranging in $[0.6, 0.8]$.

$n_{T_1} > n_{T_2}$	
T_1	T_2
$[Q^{T_1}_{2.5\%}, Q^{T_1}_{50\%}, Q^{T_1}_{97.5\%}]$	$[Q^{T_2}_{2.5\%}, Q^{T_2}_{50\%}, Q^{T_2}_{97.5\%}]$
$[0.0850, 0.1114, 0.1360]$	$[0.0767, 0.1047, 0.1293]$
$[0.1069, 0.1371, 0.1654]$	$[0.0805, 0.1127, 0.1409]$
$[0.1471, 0.1807, 0.2121]$	$[0.0925, 0.1191, 0.1453]$
$[0.0665, 0.0953, 0.1200]$	$[0.0577, 0.0824, 0.1046]$
$[0.0410, 0.0688, 0.0930]$	$[0.0018, 0.0283, 0.0486]$
$[0.0754, 0.1028, 0.1280]$	$[0.0572, 0.0827, 0.1062]$
$[0.1030, 0.1310, 0.1582]$	$[0.0701, 0.0981, 0.1224]$
$[0.1489, 0.1803, 0.2147]$	$[0.0524, 0.0836, 0.1084]$
$[0.0603, 0.0882, 0.1118]$	$[0.0295, 0.0549, 0.0773]$
$[0.1315, 0.1657, 0.1960]$	$[0.0.0772, 0.1068, 1319]$
$[0.0625, 0.0873, 0.1111]$	$[0.0490, 0.0767, 0.1000]$

TABLE K.2: In this table we report the set of quantiles obtained from the distribution of the multifractality proxy $\Delta H(1,2)$ on the two tranches T_1 and T_2 for 1000 simulated time series whose hierarchical order $n_{T_1} > n_{T_2}$. Probabilities p's of the model are initialised to take values in $[0.6, 0.8]$. One observes an almost rigid shift towards smaller values of the distribution of the observed $\Delta H(1,2)$ when the order decreases.

$n_{T_1} = n_{T_2}$	
T_1	T_2
$[Q^{T_1}_{2.5\%}, Q^{T_1}_{50\%}, Q^{T_1}_{97.5\%}]$	$[Q^{T_2}_{2.5\%}, Q^{T_2}_{50\%}, Q^{T_2}_{97.5\%}]$
$[0.1055, 0.1377, 0.1640]$	$[0.1120, 0.1457, 0.1722]$
$[0.1139, 0.1452, 0.1728]$	$[0.1254, 0.1586, 0.1893]$
$[0.0818, 0.1093, 0.1376]$	$[0.0844, 0.1124, 0.1391]$
$[0.0606, 0.0883, 0.1111]$	$[0.0595, 0.0875, 0.1118]$

TABLE K.3: In this table we report the set of quantiles obtained from the distribution of the multifractality proxy $\Delta H(1,2)$ on the two tranches T_1 and T_2 for 1000 simulated time series whose hierarchical order $n_{T_1} = n_{T_2}$. Probabilities p's of the model are initialised to take values in $[0.6, 0.8]$.

$n_{T_1} < n_{T_2}$	
T_1	T_2
$[Q^{T_1}_{2.5\%}, Q^{T_1}_{50\%}, Q^{T_1}_{97.5\%}]$	$[Q^{T_2}_{2.5\%}, Q^{T_2}_{50\%}, Q^{T_2}_{97.5\%}]$
$[0.0908, 0.1224, 0.1534]$	$[0.1787, 0.2230, 0.2714]$
$[0.0270, 0.0527, 0.0760]$	$[0.1153, 0.1509, 0.1862]$
$[0.1167, 0.1509, 0.1893]$	$[0.1661, 0.2023, 0.2440]$
$[0.1175, 0.1519, 0.1861]$	$[0.1590, 0.2004, 0.2519]$
$[0.1138, 0.1502, 0.1866]$	$[0.1808, 0.2231, 0.2705]$
$[0.1172, 0.1508, 0.1883]$	$[0.1394, 0.1773, 0.2242]$
$[0.0603, 0.0908, 0.1172]$	$[0.0905, 0.1223, 0.1540]$
$[0.0902, 0.1233, 0.1535]$	$[0.1613, 0.2018, 0.2477]$
$[0.1170, 0.1541, 0.1900]$	$[0.1408, 0.1772, 0.2189]$
$[0.1189, 0.1536, 0.1893]$	$[0.1820, 0.2225, 0.2699]$

TABLE K.4: In this table we report the set of quantiles obtained from the distribution of the multifractality proxy $\Delta H(1,2)$ on the two tranches T_1 and T_2 for 1000 simulated time series whose hierarchical order $n_{T_1} < n_{T_2}$. Probabilities p's of the model are initialised to take values in $[0.2, 0.4]$.

$n_{T_1} > n_{T_2}$	
T_1	T_2
$[Q^{T_1}_{2.5\%}, Q^{T_1}_{50\%}, Q^{T_1}_{97.5\%}]$	$[Q^{T_2}_{2.5\%}, Q^{T_2}_{50\%}, Q^{T_2}_{97.5\%}]$
$[0.1419, 0.1787, 0.2168]$	$[0.1179, 0.1525, 0.1889]$
$[0.1642, 0.2034, 0.2455]$	$[0.1403, 0.1775, 0.2193]$
$[0.2083, 0.2584, 0.3103]$	$[0.1407, 0.1772, 0.2172]$
$[0.1195, 0.1532, 0.1856]$	$[0.0886, 0.1240, 0.1537]$
$[0.0904, 0.1232, 0.1548]$	$[0.0285, 0.0536, 0.0759]$
$[0.1162, 0.1507, 0.1880]$	$[0.0926, 0.1236, 0.1540]$
$[0.1624, 0.2006, 0.2456]$	$[0.1128, 0.1509, 0.1922]$
$[0.2127, 0.2571, 0.3112]$	$[0.0899, 0.1235, 0.1556]$
$[0.1171, 0.1514, 0.1901]$	$[0.0600, 0.0888, 0.1162]$
$[0.1938, 0.2401, 0.2904]$	$[0.1180, 0.1534, 0.1911]$
$[0.1147, 0.1516, 0.1906]$	$[0.0891, 0.1214, 0.1543]$

TABLE K.5: In this table we report the set of quantiles obtained from the distribution of the multifractality proxy $\Delta H(1,2)$ on the two tranches T_1 and T_2 for 1000 simulated time series whose hierarchical order $n_{T_1} > n_{T_2}$. Probabilities p's of the model are initialised to take values in $[0.2, 0.4]$. One observes an almost rigid shift towards smaller values of the distribution of the observed $\Delta H(1,2)$ when the order decreases.

$n_{T_1} = n_{T_2}$	
T_1	T_2
$[Q^{T_1}_{2.5\%}, Q^{T_1}_{50\%}, Q^{T_1}_{97.5\%}]$	$[Q^{T_2}_{2.5\%}, Q^{T_2}_{50\%}, Q^{T_2}_{97.5\%}]$
$[0.1646, 0.2026, 0.2478]$	$[0.1641, 0.2016, 0.2472]$
$[0.1795, 0.2226, 0.2728]$	$[0.1824, 0.2226, 0.2671]$
$[0.1427, 0.1790, 0.2191]$	$[0.1393, 0.1778, 0.2153]$
$[0.0927, 0.1237, 0.1548]$	$[0.0924, 0.1228, 0.1551]$

TABLE K.6: In this table we report the set of quantiles obtained from the distribution of the multifractality proxy $\Delta H(1,2)$ on the two tranches T_1 and T_2 for 1000 simulated time series whose hierarchical order $n_{T_1} = n_{T_2}$. Probabilities p's of the model are initialised to take values in $[0.2, 0.4]$.

Bibliography

[1] R.N. Mantegna and H.E. Stanley. *An introduction to econophysics: correlations and complexity in finance*. Cambridge Univ Pr, 2000.

[2] M.M. Dacorogna. *An introduction to high-frequency finance*. Academic Pr, 2001.

[3] Anirban Chakraborti, Ioane Muni Toke, Marco Patriarca, and Frédéric Abergel. Econophysics review: I. empirical facts. *Quantitative Finance*, 11(7):991–1012, 2011.

[4] Jean-Philippe Bouchaud. The (unfortunate) complexity of the economy. *arXiv preprint arXiv:0904.0805*, 2009.

[5] Damien Challet, Matteo Marsili, and Yi-Cheng Zhang. *Minority games: interacting agents in financial markets*. Oxford University Press, 2011.

[6] Benoit B Mandelbrot and Richard L Hudson. *The Misbehavior of Markets: A fractal view of financial turbulence*. Basic books, 2007.

[7] Thomas Lux and Michele Marchesi. Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature*, 397(6719):498–500, 1999.

[8] R.N. Mantegna and H.E. Stanley. Scaling behaviour in the dynamics of an economic index. *Nature*, 376(6535):46–49, 1995.

[9] Nassim Taleb. *Fooled by randomness: The hidden role of chance in life and in the markets*. Random House Trade Paperbacks, 2005.

[10] N.N.Taleb. *The Black Swan: The Impact of the Highly Improbable.* The Random House Publishing Group, 2007.

[11] Peter Bernstein. *Against the Gods: the remarkable story of risk.* Wiley New York, 1996.

[12] Didier Sornette. *Why stock markets crash: critical events in complex financial systems.* Princeton University Press, 2009.

[13] T. Di Matteo. Multi-scaling in finance. *Quantitative Finance*, 7(1):21–36, 2007.

[14] R.N. Mantegna. Hierarchical structure in financial markets. *The European Physical Journal B-Condensed Matter and Complex Systems*, 11(1):193–197, 1999.

[15] J.W. Kantelhardt, S.A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, and H.E. Stanley. Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A*, 316(1):87–114, 2002.

[16] W.-X. Zhou. The components of empirical multifractality in financial returns. *EPL*, 88(2):28004, 2009.

[17] J. Barunik, T. Aste, T. Di Matteo, and R. Liu. Understanding the source of multifractality in financial markets. *Physica A*, 391:4234–4251, 2012.

[18] R. Morales, T. Di Matteo, R. Gramatica, and T. Aste. Dynamical generalized hurst exponent as a tool to monitor unstable periods in financial time series. *Physica A*, 391:3180–3189, 2012.

[19] R. Morales, T. Di Matteo, and T. Aste. Non-stationary multifractality in stock returns. *Physica A*, 392:6470–6483, 2013.

[20] E. Bacry, J. Delour, and J.-F. Muzy. Multifractal random walk. *Physical Review E*, 64(2):026103, 2001.

[21] R. Morales, T. Di Matteo, and T. Aste. Dependency structure and scaling properties of financial time series are related. *Scientific reports*, 4:4589; DOI:10.1038/srep04589, 2014.

[22] W.M. Song, T. Di Matteo, and T. Aste. Hierarchical information clustering by means of topologically embedded graphs. *PloS One*, 7(3):e31929, 2012.

[23] A. Meucci. *Risk and asset allocation*. Springer Verlag, 2009.

[24] J.-F. Muzy, D. Sornette, J. Delour, and A. Arneodo. Multifractal returns and hierarchical portfolio theory. *Quantitative Finance*, 1(1):131–148, 2001.

[25] Alan G Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90, 1971.

[26] R. Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1:223–236, 2001.

[27] J.-P. Bouchaud and M. Potters. *Theory of financial risk and derivative pricing: from statistical physics to risk management*. Cambridge Univ Pr, 2003.

[28] L. Calvet and A. Fisher. Multifractality in asset returns: theory and evidence. *Review of Economics and Statistics*, 84(3):381–406, 2002.

[29] Louis Bachelier. *Théorie de la spéculation*. Gauthier-Villars, 1900.

[30] Johannes Voit. *The statistical mechanics of financial markets*, volume 2. Springer Berlin, 2003.

[31] Neil F Johnson, Paul Jefferies, and Pak Ming Hui. Financial market complexity. *OUP Catalogue*, 2003.

[32] B.B. Mandelbrot. *Fractals and scaling in finance: discontinuity, concentration, risk*. Springer Verlag, 1997.

[33] T. Preis. Econophysics: complex correlations and trend switchings in financial time series. *The European Physical Journal-Special Topics*, 194(1):5–86, 2011.

[34] B.B. Mandelbrot. The variation of certain speculative prices. *The journal of business*, 36(4):394–419, 1963.

[35] Pierre Cizeau, Yanhui Liu, Martin Meyer, C-K Peng, and H Eugene Stanley. Volatility distribution in the s&p500 stock index. *Physica A: Statistical Mechanics and its Applications*, 245(3):441–445, 1997.

[36] R. Cont. Volatility clustering in financial markets: Empirical facts and agent-based models. 2009.

[37] R. Cont. Benoit Mandelbrot et la modélisation mathématique des risques financiers. URL http://hal.archives-ouvertes.fr/docs/00/66/13/71/PDF/GazetteDesMathematiciens2012.pdf.

[38] T. Di Matteo, T. Aste, and M.M. Dacorogna. Scaling behaviors in differently developed markets. *Physica A*, 324(1):183–188, 2003.

[39] T. Di Matteo, T. Aste, and M.M. Dacorogna. Long-term memories of developed and emerging markets: Using the scaling analysis to characterize their stage of development. *Journal of Banking & Finance*, 29(4):827–851, 2005.

[40] Kaushik Matia, Yosef Ashkenazy, and H Eugene Stanley. Multifractal properties of price fluctuations of stocks and commodities. *EPL (Europhysics Letters)*, 61(3):422, 2003.

[41] J Kwapień, S Drożdż, et al. Components of multifractality in high-frequency stock returns. *Physica A: Statistical Mechanics and its Applications*, 350(2): 466–474, 2005.

[42] J.-P. Bouchaud and M. Potters. More stylized facts of financial markets: leverage effect and downside correlations. *Physica A*, 299:60–70, 2001.

[43] P. Cizeau, M. Potters, and J.-P. Bouchaud. Correlation structure of extreme stock returns. *Quantitative Finance*, 1:217–222, 2001.

[44] F. Longin and B. Solnik. Extreme correlation of international equity markets. *The Journal of Finance*, 56(2):649–676, 2001.

[45] K.J. Forbes and R. Rigobon. No contagion, only interdependence: measuring stock market comovements. *The Journal of Finance*, 57(5):2223–2261, 2002.

[46] John Y. Campbell, S. Grossmann, and J. Wang. Trading volume and serial correlation in stock returns. *Quarterly Journal of Economics*, 108:905–939, 1993.

[47] T.G. Andersen. Return volatility and trading volume: an information flow interpretation of stochastic volatility. *The Journal of Finance*, 51:169–204, 1996.

[48] Willliam Feller. *An introduction to probability theory and its applications*, volume 2. John Wiley & Sons, 2008.

[49] Parameswaran Gopikrishnan, Vasiliki Plerou, Luis A Nunes Amaral, Martin Meyer, and H Eugene Stanley. Scaling of the distribution of fluctuations of financial market indices. *Physical Review E*, 60(5):5305, 1999.

[50] Vasiliki Plerou, Parameswaran Gopikrishnan, Luis A Nunes Amaral, Martin Meyer, and H Eugene Stanley. Scaling of the distribution of price fluctuations of individual companies. *Physical Review E*, 60(6):6519, 1999.

[51] Aaron Clauset, Cosma Rohilla Shalizi, and Mark EJ Newman. Power-law distributions in empirical data. *SIAM review*, 51(4):661–703, 2009.

[52] Bruce M Hill. A simple general approach to inference about the tail of a distribution. *The annals of statistics*, pages 1163–1174, 1975.

[53] Fabio Clementi, T Di Matteo, and Mauro Gallegati. The power-law tail exponent of income distributions. *Physica A: Statistical Mechanics and its Applications*, 370(1):49–53, 2006.

[54] Eugene Fama. The behaviour of stock-market prices. *The Journal of Business*, 38(1):34–105, 1965.

[55] B.B. Mandelbrot. How long is the coast of britain? statistical self-similarity and fractional dimension. *Science*, 156(3775):636–638, 1967.

[56] B.B. Mandelbrot and Howard M Taylor. On the distribution of stock price differences. *Operations research*, 15(6):1057–1062, 1967.

[57] B.B. Mandelbrot and John W Van Ness. Fractional brownian motions, fractional noises and applications. *SIAM review*, 10(4):422–437, 1968.

[58] Shoaleh Ghashghaie, Wolfgang Breymann, Joachim Peinke, Peter Talkner, and Yadollah Dodge. Turbulent cascades in foreign exchange markets. *Nature*, 381(6585):767–770, 1996.

[59] P.Sollich et al., editor. *Scaling and Multi-Scaling in Financial Markets*, volume 553, pages 297–302. Disordered Systems and Complex Systems, AIP Conference Proceedings, 2001.

[60] L. Calvet and A. Fisher. Forecasting multifractal volatility. *Journal of econometrics*, 105(1):27–58, 2001.

[61] R. Liu, T.D. Matteo, and T. Lux. Multifractality and long-range dependence of asset returns: The scaling behaviour of the markov-switching multifractal model with lognormal volatility components. *Advances in Complex Systems*, 11(5):669–684, 2008.

[62] T. Lux. The markov-switching multifractal model of asset returns. *Journal of business & economic statistics*, 26(2):194–210, 2008.

[63] R. Liu, T. Di Matteo, and T. Lux. True and apparent scaling: The proximity of the markov-switching multifractal model to long-range dependence. *Physica A*, 383(1):35–42, 2007.

[64] M. Bartolozzi, C. Mellen, T. Di Matteo, and T. Aste. Multi-scale correlations in different futures markets. *The European Physical Journal B*, 58(2):207–220, 2007.

[65] Ladislav Kristoufek. Fractal markets hypothesis and the global financial crisis: scaling, investment horizons and liquidity. *Advances in Complex Systems*, 15(06), 2012.

[66] Z.-Q.-Q. Jiang and W.-X. Zhou. Multifractality in stock indexes: Fact or fiction? *Physica A: Statistical Mechanics and its Applications*, 387(14): 3605–3614, 2008.

[67] B.B. Mandelbrot, A. Fisher, and L. Calvet. The multifractal model of asset returns. *Cowles Foundation discussion paper no. 1164, Yale University*, 1997.

[68] J.-P. Bouchaud, M. Potters, and M. Meyer. Apparent multifractality in financial time series. *The European Physical Journal B*, 13(3):595–599, 2000.

[69] Z. Ding, C.W.J. Granger, and R.F. Engle. A long memory property of stock market returns and a new model. *Journal of empirical finance*, 1(1):83–106, 1993.

[70] Thomas Lux. Turbulence in financial markets: the surprising explanatory power of simple cascade models. *Quantitative finance*, 1(6):632–640, 2001.

[71] Ulrich A Müller, Michel M Dacorogna, Richard B Olsen, Olivier V Pictet, Matthias Schwarz, and Claude Morgenegg. Statistical study of foreign exchange rates, empirical evidence of a price change scaling law, and intraday analysis. *Journal of Banking & Finance*, 14(6):1189–1208, 1990.

[72] J.-F. Muzy, E. Bacry, and A. Kozhemyak. Extreme values and fat tails of multifractal fluctuations. *Physical Review E*, 73(6):066114, 2006.

[73] J.-F. Muzy, E. Bacry, R. Baile, and P. Poggi. Uncovering latent singularities from multifractal scaling laws in mixed asymptotic regime. application to turbulence. *EPL (Europhysics Letters)*, 82(6):60007, 2008.

[74] E Bacry, A Kozhemyak, and J.-F. Muzy. Log-normal continuous cascade model of asset returns: aggregation properties and estimation. *Quantitative Finance*, 13(5):795–818, 2013.

[75] J. Barunik and L. Kristoufek. On hurst exponent estimation under heavy-tailed distributions. *Physica A*, 389(18):3844–3855, 2010.

[76] P.A. Groenendijk, A. Lucas, and C.G. Vries. A hybrid joint moment ratio test for financial time series. *Preprint of Erasmus University*, (1):1–38, 1998. (http://www.few.eur.nl/few/people/cdevries/).

[77] B. LeBaron. Stochastic volatility as a simple generator of apparent financial power laws and long memory. *Quantitative Finance*, 1:621–631, 2001.

[78] A. Carbone, G. Castelli, and HE Stanley. Time-dependent hurst exponent in financial time series. *Physica A*, 344(1):267–271, 2004.

[79] D. Grech and G. Pamuła. The local hurst exponent of the financial time series in the vicinity of crashes on the polish stock exchange market. *Physica A*, 387(16):4299–4308, 2008.

[80] Ladislav Kristoufek. Local scaling properties and market turning points at prague stock exchange. *Acta Physica Polonica B*, 41(6):1223, 2010.

[81] Krzysztof Domino. The use of the hurst exponent to predict changes in trends on the warsaw stock exchange. *Physica A*, 390(1):98–109, 2011.

[82] Krzysztof Domino. The use of the hurst exponent to investigate the global maximum of the warsaw stock exchange wig20 index. *Physica A*, 391(1): 156–169, 2012.

[83] F. Pozzi, T. Di Matteo, and T. Aste. Exponential smoothing weighted correlations. *The European Physical Journal B*, 85:175, 2012.

[84] Thomas Helbling, Valerie Mercer-Blackman, and Kevin Cheng. Commodities boom: riding a wave. *Finance and Development, IMF*, March 2008.

[85] Thomas Helbling. Commodities in boom. *Finance and Development, IMF*, 49(2), 2012.

[86] T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327, 1986.

[87] L.E. Calvet and A. Fisher. *Multifractal Volatility: Theory, Forecasting, and Pricing*. Academic Pr, 2008.

[88] J.-F. Muzy, J. Delour, and E. Bacry. Modelling fluctuations of financial time series: from cascade process to stochastic volatility model. *The European Physical Journal B*, 17(3):537–548, 2000.

[89] Josep Perelló, Jaume Masoliver, and Jean-Philippe Bouchaud. Multiple time scales in volatility and leverage correlations: a stochastic volatility model. *Applied Mathematical Finance*, 11(1):27–50, 2004.

[90] Danilo Delpini and Giacomo Bormetti. Stochastic volatility with heterogeneous time scales. *arXiv preprint arXiv:1206.0026*, 2012.

[91] S. Miccichè, G. Bonanno, F. Lillo, and R.N. Mantegna. Volatility in financial markets: stochastic models and empirical results. *Physica A*, 314:66–73, 2002.

[92] F. Baldovin and A. Stella. Scaling and efficiency determine the irreversible evolution of a market. *Proceedings of the National Academy of Sciences of the United States of America*, 104(50):19741–19744, 2007.

[93] A. Arneodo, E. Bacry, S. Manneville, and J.-F. Muzy. Analysis of random cascades using space-scale correlation functions. *Physical Review Letters*, 80 (4):708–711, 1998.

[94] E. Bacry, A. Kozhemyak, and J.-F. Muzy. Continuous cascade models for asset returns. *Journal of Economic Dynamics and Control*, 32(1):156–199, 2008.

[95] J.-F. Muzy, R. Baïle, and P. Poggi. Intermittency of surface-layer wind velocity series in the mesoscale range. *Physical Review E*, 81(5):056308, 2010.

[96] L. Borland, J.-P. Bouchaud, J.-F. Muzy, and G. Zumbach. The dynamics of financial marketsmandelbrots multifractal cascades. *Wilmott Magazine*, page p 86, 2005.

[97] L. Borland and J.-P. Bouchaud. On a multi-timescale statistical feedback model for volatility fluctuations. *arXiv:physics.soc-ph/0507073*, 2005.

[98] James Douglas Hamilton. *Time series analysis*, volume 2. Cambridge Univ Press, 1994.

[99] M. Tumminello, T. Aste, T. Di Matteo, and R.N. Mantegna. A tool for filtering information in complex systems. *Proceedings of the National Academy of Sciences of the United States of America*, 102(30):10421, 2005.

[100] Michael R Anderberg. Cluster analysis for applications. Technical report, DTIC Document, 1973.

[101] Mark S Aldenderfer and Roger K Blashfield. Cluster analysis: Quantitative applications in the social sciences. *Beverly Hills: Sage Publication*, 1984.

[102] G. Caldarelli. *Scale-Free Networks: Complex webs in nature and technology*. Oxford University Press, USA, 2007.

[103] T. Di Matteo, T. Aste, S.T. Hyde, and Ramsden S. Interest rates hierarchical structure. *Physica A*, 355:21–33, 2005.

[104] T. Aste, T. Di Matteo, and S.T. Hyde. Complex networks on hyperbolic surfaces. *Physica A*, 346:20–26, 2005.

[105] T. Aste, R. Gramatica, and T. Di Matteo. Random and frozen states in complex triangulations. *Philosophical Magazine*, 92(1-3):244–254, 2011.

[106] T. Aste, R. Gramatica, and T. Di Matteo. Exploring complex networks via topological embeddings on surfaces. *Physical Review E*, 86:036109, 2012.

[107] Giovanni Bonanno, Nicolas Vandewalle, and Rosario N Mantegna. Taxonomy of stock market indices. *Physical Review E*, 62(6):R7615, 2000.

[108] G. Bonanno, G. Caldarelli, F. Lillo, S. Miccichè, N. Vandewalle, and R.N. Mantegna. Networks of equities in financial markets. *The European Physical Journal B-Condensed Matter and Complex Systems*, 38(2):363–371, 2004.

[109] V. Tola, F. Lillo, M. Gallegati, and R.N. Mantegna. Cluster analysis for portfolio optimization. *Journal of Economic Dynamics and Control*, 32(1): 235–258, 2008.

[110] M. Tumminello, F. Lillo, and R.N. Mantegna. Hierarchically nested factor model from multivariate data. *EPL (Europhysics Letters)*, 78:30006, 2007.

[111] C Coronnello, M Tumminello, FABRIZIO Lillo, S Micciche, and ROSARIO NUNZIO Mantegna. Sector identification in a set of stock return time series traded at the london stock exchange. *arXiv preprint cond-mat/0508122*, 2005.

[112] John C Gower and GJS Ross. Minimum spanning trees and single linkage cluster analysis. *Applied statistics*, pages 54–64, 1969.

[113] W.M. Song. *Understanding Complex Systems: Network Approach to Information Filtering*. PhD thesis, Australian National University, 2012.

[114] Michele Tumminello, Salvatore Miccichè, Fabrizio Lillo, Jan Varho, Jyrki Piilo, and Rosario N Mantegna. Community characterization of heterogeneous complex systems. *Journal of Statistical Mechanics: Theory and Experiment*, 2011(01):P01019, 2011.

[115] N. Musmeci and al. Dynamical properties of clusters in nyse. in preparation, 2013.

[116] Bradley Efron. Nonparametric estimates of standard error: The jackknife, the bootstrap and other methods. *Biometrika*, 68(3):589–599, 1981.

[117] Ulrich A Müller, Roland Bürgi, and Michel M Dacorogna. Bootstrapping the economy–a non-parametric method of generating consistent future scenarios. *MPRA Paper, University Library of Munich*, 2004.

[118] Christian Borghesi, Matteo Marsili, and Salvatore Micciche. Emergence of time-horizon invariant correlation structure in financial returns by subtraction of the market mode. *Physical Review E*, 76(2):026104, 2007.

[119] J-P Onnela, Anirban Chakraborti, Kimmo Kaski, and Janos Kertesz. Dynamic asset trees and black monday. *Physica A*, 324(1):247–252, 2003.

[120] T. Aste, W. Shaw, and T. Di Matteo. Correlation structure and dynamics in volatile markets. *New Journal of Physics*, 12:085009, 2010.

[121] J-P Onnela, Anirban Chakraborti, Kimmo Kaski, Janos Kertesz, and Antti Kanto. Dynamics of market correlations: Taxonomy and portfolio analysis. *Physical Review E*, 68(5):056110, 2003.

[122] S. Miccichè. Empirical study of the relationship between the cross-correlation among stocks and stocks'volatility clustering. *Journal of Statistical Mechanics*, page P0515, 2013.

[123] Christian M Hafner and Hans Manner. Multivariate time series models for asset prices. In *Handbook of Computational Finance*, pages 89–115. Springer, 2012.

[124] F. Lindskog and A.J. McNeil. Common poisson shock models: applications to insurance and credit risk modelling. *Astin Bulletin*, 33(2):209–238, 2003.

[125] Harry Markowitz. Portfolio selection. *The journal of finance*, 7(1):77–91, 1952.

[126] Harry Markowitz. *Portfolio selection: efficient diversification of investments*. Yale university press, 1959.

[127] P. Embrechts, A. McNeil, and D. Straumann. Correlation and dependence in risk management: properties and pitfalls. *Risk management: value at risk and beyond*, pages 176–223, 2002.

[128] Giulia Iori and Saqib Jafarey. Criticality in a model of banking crises. *Physica A: Statistical Mechanics and its Applications*, 299(1):205–212, 2001.

[129] Giulia Iori, Saqib Jafarey, and Francisco G Padilla. Systemic risk on the interbank market. *Journal of Economic Behavior & Organization*, 61(4): 525–542, 2006.

[130] Frank Schweitzer, Giorgio Fagiolo, Didier Sornette, Fernando Vega-Redondo, and Douglas R White. Economic networks: What do we know and what do we need to know? *Advances in Complex Systems*, 12(04n05):407–422, 2009.

[131] Giulia Iori, Saqib Jafarey, and Francisco Padilla. Interbank lending and systemic risk. *Journal of Economic Behaviour and Organisation*, 2003.

[132] A.J. McNeil, R. Frey, and P. Embrechts. *Quantitative risk management: Concepts, techniques and tools*. Princeton Univ Pr, 2005.

[133] D.X. Li. On default correlation: a copula function approach. *Journal of Fixed income*, 9(4):43–54, 2000.

[134] Damiano Brigo, Andrea Pallavicini, and Roberto Torresetti. *Credit models and the crisis: A journey into CDOs, copulas, correlations and dynamic models*. Wiley New York, 2010.

[135] L. Bauwens, S. Laurent, and J.V.K. Rombouts. Multivariate garch models: a survey. *Journal of Applied Econometrics*, 21(1):79–109, 2006.

[136] Manabu Asai, Michael McAleer, and Jun Yu. Multivariate stochastic volatility: a review. *Econometric Reviews*, 25(2-3):145–175, 2006.

[137] Rémy Chicheportiche and Jean-Philippe Bouchaud. The fine-structure of volatility feedback. *arXiv preprint arXiv:1206.2153*, 2012.

[138] John Y. Campbell and Andrew W. Lo. *The econometrics of financial markets*. Princeton University Press, 1997.

[139] Gregory Connor. The three types of factor models: a comparison of their explanatory power. *Financial Analysts Journal*, pages 42–46, 1995.

[140] Kai-Tai Fang and Yao-Ting Zhang. *Generalized multivariate analysis*. Springer, 1990.

[141] Kai-Tai Fang, Samuel Kotz, and Kai Wang Ng. *Symmetric Multivariate and Related Distributions Monographs on Statistics and Applied Probability*. London: Chapman and Hall Ltd. MR1071174, 1990.

[142] Rémy Chicheportiche and Jean-Philippe Bouchaud. The joint distribution of stock returns is not elliptical. *International Journal of Theoretical and Applied Finance*, 15(03), 2012.

[143] Rémy Chicheportiche. *Non-linear Dependence in Finance*. PhD thesis, Ecole Centrale Paris, 2013.

[144] Y. Malevergne and D. Sornette. Testing the gaussian copula hypothesis for financial assets dependences. *Quantitative Finance*, 3(4):231–250, 2003.

[145] Kanti V Mardia. Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57(3):519–530, 1970.

[146] Kanti V Mardia. Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies. *Sankhyā: The Indian Journal of Statistics, Series B*, pages 115–128, 1974.

[147] Y. Malevergne and D. Sornette. *Extreme financial risks: From dependence to risk management*. Springer Verlag, 2006.

[148] Melvin Dale Springer. *The algebra of random variables*. Wiley New York, 1979.

[149] Stefano Demarta and Alexander J McNeil. The t copula and related copulas. *International Statistical Review*, 73(1):111–129, 2005.

[150] T. Bollerslev. Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch approach. *Review of Economics and Statistics*, 72:498–505, 1990.

[151] M.C. Münnix, T. Shimada, R. Schäfer, F.L.T.H. Seligman, T. Guhr, and HE Stanley. Identifying states of a financial market. *Arxiv preprint arXiv:1202.1623*, 2012.

[152] Bence Tóth and János Kertész. Increasing market efficiency: Evolution of cross-correlations of stock returns. *Physica A*, 360(2):505–515, 2006.

[153] J Shen and B Zheng. Cross-correlation in financial dynamics. *EPL (Europhysics Letters)*, 86(4):48005, 2009.

[154] Thomas Conlon, Heather J Ruskin, and Martin Crane. Cross-correlation dynamics in financial time series. *Physica A: Statistical Mechanics and its Applications*, 388(5):705–714, 2009.

[155] Robert Engle. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3):339–350, 2002.

[156] Yiu K Tse and Albert K C Tsui. A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *Journal of Business & Economic Statistics*, 20(3):351–362, 2002.

[157] Christian M Hafner and Philip Hans Franses. A generalized dynamic conditional correlation model: simulation and application to many assets. *Econometric Reviews*, 28(6):612–631, 2009.

[158] Tim Bollerslev, Robert F Engle, and Jeffrey M Wooldridge. A capital asset pricing model with time-varying covariances. *The Journal of Political Economy*, pages 116–131, 1988.

[159] Robert F Engle and Kenneth F Kroner. Multivariate simultaneous generalized arch. *Econometric theory*, 11(01):122–150, 1995.

[160] Andrew Harvey, Esther Ruiz, and Neil Shephard. Multivariate stochastic variance models. *The Review of Economic Studies*, 61(2):247–264, 1994.

[161] Manabu Asai and Michael McAleer. Asymmetric multivariate stochastic volatility. *Econometric Reviews*, 25(2-3):453–473, 2006.

[162] Jon Danielsson. Stochastic volatility in asset prices estimation with simulated maximum likelihood. *Journal of Econometrics*, 64(1):375–400, 1994.

[163] Lillo F. and R.N. Mantegna. Spectral density of the correlation matrix of factor model: a random matrix theory approach. *Physical Review E*, 72: 016219, 2005.

[164] Matteo Marsili. Dissecting financial markets: sectors and states. *Quantitative Finance*, 2(4):297–302, 2002.

[165] Jun Yu and Renate Meyer. Multivariate stochastic volatility models: Bayesian estimation and model comparison. *Econometric Reviews*, 25(2-3): 361–384, 2006.

[166] José M Quintana and Mike West. An analysis of international exchange rates using multivariate dlm's. *The Statistician*, pages 275–281, 1987.

[167] E. Bacry, J Delour, and J.-F. Muzy. A multivariate multifractal model for return fluctuations. *arXiv preprint cond-mat/0009260*, 2000.

[168] D. Brigo, A. Pallavicini, and R. Torresetti. Cluster-based extension of the generalized poisson loss dynamics and consistency with single names. *International Journal of Theoretical and Applied Finance*, 10(4):607–631, 2007.

[169] Daryl J Daley and David Vere-Jones. *An introduction to the theory of point processes: volume II: general theory and structure*, volume 2. Springer, 2007.

[170] Yosihiko Ogata. Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association*, 83(401):9–27, 1988.

[171] Yosihiko Ogata. Seismicity analysis through point-process modeling: A review. *Pure and Applied Geophysics*, 155(2-4):471–507, 1999.

[172] Paul Embrechts, Thomas Liniger, and Lu Lin. Multivariate hawkes processes: an application to financial data. *Journal of Applied Probability*, 48: 367–378, 2011.

[173] V. Chavez-Demoulin, A.C. Davison, and A.J. McNeil. Estimating value-at-risk: a point process approach. *Quantitative Finance*, 5(2):227–234, 2005.

[174] E. Bacry, Sylvain Delattre, Marc Hoffmann, and J.-F. Muzy. Modelling microstructure noise with mutually exciting point processes. *Quantitative Finance*, 13(1):65–77, 2013.

[175] Clive G Bowsher. Modelling security market events in continuous time: Intensity based, multivariate point process models. *Journal of Econometrics*, 141(2):876–912, 2007.

[176] Luc Bauwens and Nikolaus Hautsch. *Modelling financial high frequency data using point processes*. Springer, 2009.

[177] Yacine Aït-Sahalia, Julio Cacho-Diaz, and Roger JA Laeven. Modeling financial contagion using mutually exciting jump processes. Technical report, National Bureau of Economic Research, 2010.

[178] Jaume Masoliver and Josep Perelló. Multiple time scales and the exponential ornstein–uhlenbeck stochastic volatility model. *Quantitative Finance*, 6(5): 423–433, 2006.

[179] Steven L Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of financial studies*, 6(2):327–343, 1993.

[180] Rémy Chicheportiche and Jean-Philippe Bouchaud. A nested factor model for non-linear dependences in stock returns. *arXiv preprint arXiv:1309.3102*, 2013.

[181] Filip Lindskog, Alexander McNeil, and Uwe Schmock. *Kendalls tau for elliptical distributions*, volume 111, chapter Credit Risk-Measurement, Evaluation and Management, page 157. Physica Verlag, 2003.

[182] F Pozzi, T Di Matteo, and T Aste. Spread of risk across financial markets: better to invest in the peripheries. *Scientific reports*, 3, 2013.

[183] L. Calvet and A. Fisher. How to forecast long-run volatility: regime-switching and the estimation of multifractal processes. *Journal of Financial Econometrics*, 2(1):49–83, 2004.

[184] Karl Pearson. Notes on the history of correlation. *Biometrika*, 13(1):25–45, 1920.

[185] Karl Pearson. The fundamental problem of practical statistics. *Biometrika*, 13(1):1–16, 1920.

[186] Robert Litterman and Kurt Winkelmann. Estimating covariance matrices. *Risk Management Series, Goldman Sachs*, 2, 1998.

[187] Stamatis Cambanis, Steel Huang, and Gordon Simons. On the theory of elliptically contoured distributions. *Journal of Multivariate Analysis*, 11(3): 368–385, 1981.